



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

***Metabolic engineering der Taxolbiosynthese in *Saccharomyces cerevisiae****

Vom Fachbereich Chemie  
der Technischen Universität Darmstadt  
zur Erlangung des akademischen Grades eines  
Doctor rerum naturalium (Dr. rer. nat)

genehmigte

**Dissertation**

eingereicht von

Dipl.-Biologin Pia Dahm  
aus Münster

Referent:	Prof. Dr. W.-D. Fessner
Korreferent:	Prof. Dr. H. Kolmar
Tag der Einreichung:	14. Oktober 2010
Tag der mündlichen Prüfung:	17. Dezember 2010

Darmstadt 2011

D17

---

Diese Arbeit wurde in der Zeit vom September 2005 bis September 2010 unter der Leitung von Prof. Dr. W.-D. Fessner angefertigt. Die praktischen Arbeiten erfolgten extern am Fraunhofer Institut für Molekularbiologie und Pflanzenbiotechnologie in Aachen. Die Betreuung vor Ort übernahm dabei Dr. S. Jennewein.

Teile dieser Arbeit wurden bereits veröffentlicht oder auf Tagungen vorgestellt:

### **Publikationen**

„Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production”, *Metabolic engineering* 10 (2008) 201-206

“Introduction of the early pathway to Taxol biosynthesis in yeast by means of biosynthetic gene cluster construction”, *Methods in molecular biology* (2010) 643: 145-163

“Development of carbon plasma-coated multiwell plates for high-throughput mass spectrometric analysis of highly lipophilic fermentation products”, *Analytical Biochemistry* 403 (2010) 108-113

### **Poster**

“Metabolic engineering of Taxol biosynthesis in *Saccharomyces cerevisiae*”  
14<sup>th</sup> European congress in biotechnology, 13.-16. Sept. 2009, Barcelona, Spanien

# Inhalt

I	Einleitung.....	1
I.1	Terpene .....	2
I.1.1	Menthol .....	3
I.1.2	Artemisinin.....	4
I.1.3	Taxol.....	5
I.2	Biosynthese von Terpenen.....	10
I.2.1	Biosynthese der Vorstufen .....	10
I.2.2	Biosynthese der Terpengrundgerüste .....	16
I.2.3	Biosynthese von Taxol .....	17
I.3	Metabolic engineering .....	26
I.3.1	Metabolic engineering von Mikroorganismen zur Produktion von Terpenen .....	27
I.3.2	Metabolic engineering zur Produktion von Taxol.....	32
I.4	Zielsetzung.....	35
II	Material und Methoden.....	38
II.1	Mikroorganismen.....	38
II.1.1	Verwendete Stämme, Vektoren und Oligos .....	38
II.2	Nährmedien .....	44
II.2.1	2YT-Medium.....	44
II.2.2	SOC-Medium .....	45
II.2.3	YPD-Medium .....	45
II.2.4	SC-Medium .....	45
II.2.5	Medium 88 (DSM).....	45
II.2.6	Antibiotika.....	46
II.3	Handhabung von Bakterien- und Hefestämmen.....	46
II.3.1	Zellanzucht und Zellernte von Bakterienstämmen.....	46
II.3.2	Zellanzucht und Zellernte von Hefestämmen .....	46
II.3.3	Fermentation von Hefezellen im 30l-Maßstab.....	47

II.3.4	Stammkonservierung Reinheitskontrolle .....	47
II.4	Molekularbiologische Methoden .....	47
II.4.1	Isolierung von Nukleinsäuren .....	47
II.4.2	In vitro-Amplifikation von DNS-Fragmente.....	50
II.4.3	Gelelektrophoretische Auftrennung von DNS .....	53
II.4.4	Aufreinigung von DNS .....	54
II.4.5	Enzymatische Modifikation von DNS .....	55
II.4.6	In vitro -Rekombination .....	55
II.4.7	Transfer von DNS .....	61
II.5	Biochemische Methoden .....	63
II.5.1	Isolierung von Gesamtprotein .....	63
II.5.2	SDS-Polyacrylamidgelelektrophorese .....	64
II.5.3	Westernblot .....	65
II.5.4	Northernblot .....	67
II.5.5	Massenspektrometrische Analyse .....	70
II.6	Verwendete Programme .....	72
II.7	Chemikalien und Verbrauchsmaterial .....	72
II.7.1	Chemikalien und Enzyme .....	72
II.7.2	Verbrauchsmaterialien .....	72
III	Resultate.....	73
III.1	Prinzipielle Überlegungen .....	73
III.1.1	Expressionssystem und Wahl des Stammes.....	73
III.1.2	Selektion der Expressionsvektoren .....	75
III.2	Generierung eines Taxadien produzierenden Hefestamms.....	77
III.3	Generierung eines Genclusters .....	89
III.3.1	Klonierung.....	89
III.3.2	Stammanalyse.....	94
III.3.3	Modifizierung des Genclusters.....	98

III.3.4	Stabilitätsanalyse transformierter Fusionskonstrukte.....	112
III.3.5	Klonierung des Genclusters in pVV200.....	121
III.4	Konstruktion eines künstlichen Hefechromosoms.....	124
III.4.1	Modifikation von pYAC4 .....	124
III.4.2	Integration des Genclusters und Generierung eines rekombinanten Hefestamms .....	126
IV	Diskussion.....	128
IV.1	Generierung eines Taxadien produzierenden Hefestammes.....	128
IV.2	Generierung eines Genclusters .....	135
IV.2.1	Fusionierung von DNS-Fragmenten .....	136
IV.2.2	Analyse generierter Fusionsstämme.....	137
IV.2.3	Instabilität der Fusionsvektoren in der Hefe .....	140
IV.2.4	Umklonierung des Clusters in pVV200 .....	147
IV.2.5	Expression des Clusters über ein künstliches Hefechromosom .....	149
V	Zusammenfassung .....	153
VI	Referenzen .....	157
VII	Anhang.....	168

## I Einleitung

Was ist das für ein Nadelbaum, der keine Zapfen trägt, sondern feuerrote Beeren? Der in allen Teilen giftig ist außer dem Fruchtfleisch – und trotzdem von Wild- und Weidetieren verbissen wird? Der vom Küstengebiet bis in den Himalaya, von Schottland bis nach Mexiko gedeihen kann? Über den etliche Mythen und Sagen kursieren und der in allen Weltreligionen eine wichtige Stellung einnimmt? Der nicht nur zur Anfertigung von Musikinstrumenten und Handwerk diene, sondern aus dessen Holz Bögen gemacht wurden, die ganze Kriege zu entscheiden schienen und sowohl Robin Hood für die Rächung der Armen verwendete, als auch Wilhelm Tell bei der Befreiung des Schweizer Volkes diene? Der imstande ist, sich neu zu gebären und so zu ewigem Leben fähig ist? Der daher als Lebensbaum die Stellung eines Hüters der Schwelle zwischen Leben und Tod einnimmt, sodass sogar der Evolutionsforscher Charles Darwin darunter begraben werden wollte? Der von Naturvölkern als Allheilmittel eingesetzt wird, aber gleichzeitig auch zu zahlreichen Vergiftungen führt?

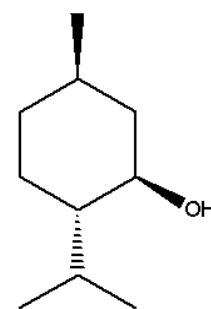


Abbildung I.1: Aus Eibenholz angefertigte Gegenstände (Quelle: Die Eibe in neuem Licht, von Fred Hageneder)

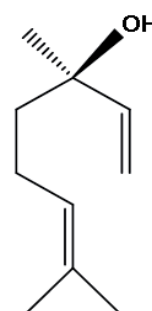
Die Eibe *Taxus* ist nicht nur kulturgeschichtlich interessant, sondern verkörpert auch in der heutigen Zeit einen wichtigen Baum. So wurde in seiner Rinde der biologisch aktive Naturstoff Taxol gefunden, der eines der meist eingesetzten Krebsmedikamente darstellt und der großen Gruppe der Terpene angehört...

## I.1 Terpene

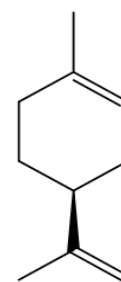
Terpene bilden mit mehr als 55000 isolierten Substanzen nicht nur die strukturell vielfältigste, sondern auch die größte Gruppe von Naturstoffen.<sup>[1]</sup> Ursprünglich leitet sich der Begriff Terpen von dem Kiefernharz Terpentin (Balsamum Terebinthinae) ab, welches neben den Harzsäuren einige herkunftsgemäß als Terpene bezeichnete Kohlenwasserstoffe enthält.<sup>[2]</sup> Die Einteilung von Terpenen basiert auf dem einheitlichen Bauprinzip aus Isopreneinheiten.<sup>[3]</sup> Je nach Anzahl dieser  $C_5$ -Untereinheiten wird zwischen Hemi- ( $C_5$ ), Mono- ( $C_{10}$ ), Sesqui- ( $C_{15}$ ), Di- ( $C_{20}$ ), Sester- ( $C_{25}$ ), Tri- ( $C_{30}$ ), Tetra- ( $C_{40}$ ) und Polyterpenen ( $C_5$ )<sub>n</sub> mit  $n > 8$  unterschieden. Als pflanzliche Sekundärmetaboliten dienen Terpene nicht dem Energie- oder anabolen, bzw. katabolen Stoffwechsel, sondern werden für kommunikative Zwecke oder den Schutz vor Fraßfeinden und Pathogenen eingesetzt.<sup>[4]</sup> Aufgrund ihrer vielen Anwendungsmöglichkeiten in weiten Bereichen industrieller Produkte, sind sie aber nicht nur in der Natur von Bedeutung, sondern auch von kommerziellem Interesse. So werden z.B. Menthol, Linalool, Limonen und verschiedene Carotinoide als Geschmack-, Duft- und Farbstoffe in Lebensmittel- bzw. Pharmaindustrie eingesetzt.<sup>[5]</sup> Durch ihre Wirkung auf Tumorzellen (Taxol), Malaria übertragende *Plasmodium* spp. (Artemisinin) und andere Mikroorganismen finden Terpene weiterhin in der Medizin Verwendung.<sup>[6]</sup> Da ihnen ebenso zugesprochen wird antioxidative und immunaktivierende Eigenschaften zu haben (Carotinoide), sind sie außerdem Inhaltstoffe sogenannter Nutrazeutika. Die Strukturformeln ausgewählter, industriell relevanter Terpene sind in Abbildung I.2 dargestellt.



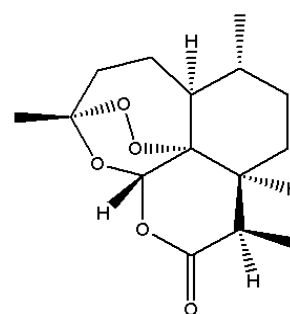
(-)-Menthol



(-)-Linalool



(-)-Limonen

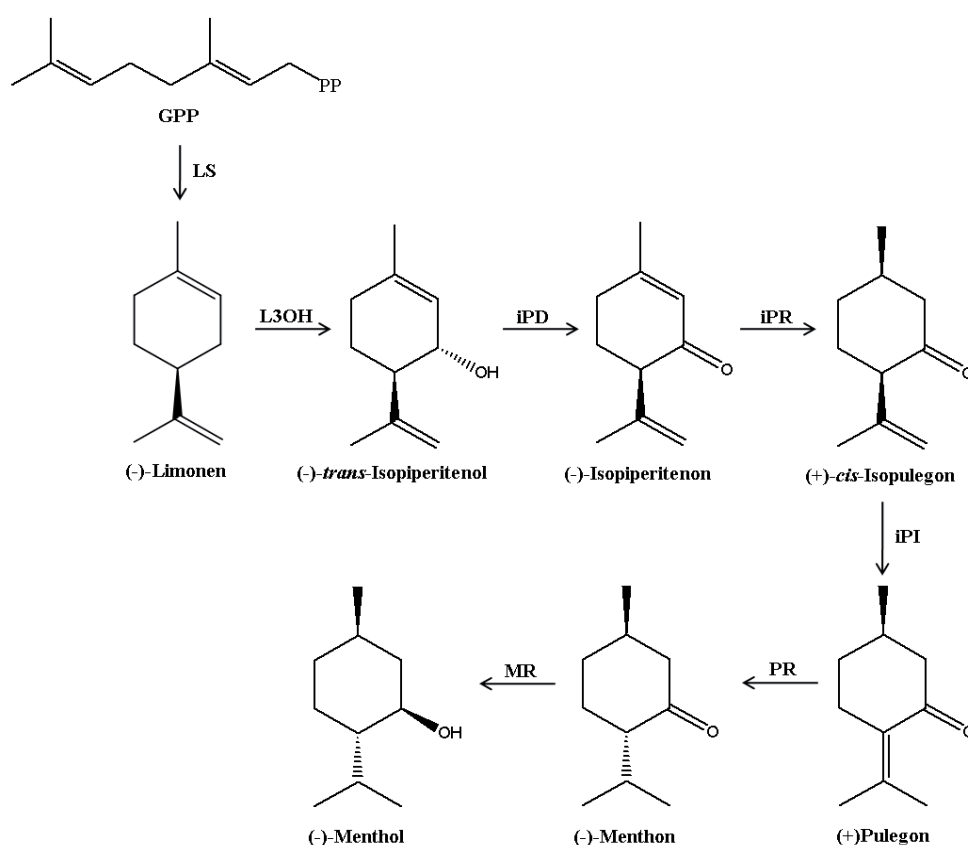


Artemisinin

Abbildung I.2: Struktur ausgesuchter Beispiele industriell relevanter Terpene

### I.1.1 *Menthol*

Menthol ist das wohl bekannteste Monoterpen. Als Hauptbestandteil des Öls von Minzen wie *Mentha piperita* wird es aufgrund seines erfrischenden Geschmacks und seiner kühlenden Wirkung in Kosmetika, Zahnpasta, Kaugummi und Zigaretten verwendet. Die antimikrobielle Wirkung macht es weiterhin zu einem beliebten Zusatz in Pharmazeutika.<sup>[7]</sup> Die Jahresproduktion von 19000 t erfolgt dabei durch chemische Synthese oder Destillation des essenziellen pflanzlichen Öls, welches einen Mentholgehalt von bis zu 90 % erreicht.<sup>[8]</sup> Wie auch andere Monoterpene wird (-)-Menthol in den hochspezialisierten sekretorischen Drüsenzellen der Pflanze produziert und akkumuliert, die in einer Anzahl von bis zu 8000 pro Blatt vorkommen.<sup>[9]</sup> Seine Biosynthese involviert mit Zyklisierung, Hydroxylierungen und Redoxtransformationen alle repräsentativen Reaktionstypen des Terpenmetabolismus (Abbildung I.3). *Mentha piperita* verkörpert daher einen Modellorganismus für Untersuchungen zur Organisation und Regulation nicht nur der (-)-Menthol-Biosynthese, sondern von Terpenen im Allgemeinen.<sup>[7]</sup>

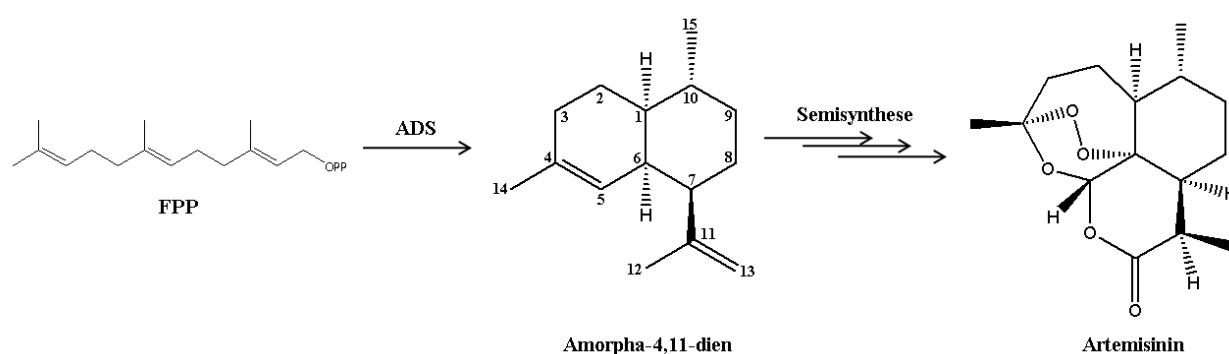


**Abbildung I.3: Schematische Darstellung des Biosynthesewegs von (-)-Menthol.** GPP: Geranyldiphosphat; LS: (-)-Limonensynthase; L3OH: (-)-Limonen-3-hydroxylase; iPD: (-)-trans-Isopiperitenoldehydrogenase; iPR: (-)-Isopiperitenonreduktase; PR: (+)-Pulegonreduktase; MR: (-)-Menthonreduktase



### I.1.2 Artemisinin

Eines der medizinisch bedeutendsten Sesquiterpene stellt das Antimalariamedikament Artemisinin dar. Es wurde aus dem Einjährigen Beifuß *Artemisia annua* isoliert, der seit Jahrhunderten in der chinesischen Medizin eingesetzt und dessen Wirkung gegen Malaria bereits seit dem 16. Jahrhundert beschrieben wird. 1967 wurde *A. annua* daher Gegenstand einer durch die Volksrepublik China initiierten Suche nach neuen medizinisch einsetzbaren Substanzen.<sup>[10]</sup> Die biologisch aktive Verbindung wurde 1972 aus *Artemisia annua* isoliert und bekam den Namen Artemisinin. Als Antimalaria-Medikament ist Artemisinin einzigartig, da es auch gegen multiresistente *Plasmodium*-Stämme wirksam ist.<sup>[11]</sup> Artemisinin-basierte Kombinationstherapien stellen daher das von der WHO empfohlene Medikament dar (*Guidelines for the treatment of malaria*). Der ursprüngliche Artemisininingehalt in *Artemisia annua* von nur 0,01-0,5 %<sup>[12]</sup> konnte zwar durch die Züchtung Artemisinin-reicher Kultivare auf 0,8-1,2 % erhöht werden, der extrahierbare Ertrag liegt allerdings bei 0,05-0,2 %, <sup>[13]</sup> sodass der weltweite Bedarf durch angebaute *Artemisia* –Spezies nicht gedeckt werden kann. Viele Ansätze zur chemischen Synthese wurden bereits gezeigt und zum Teil auch patentiert.<sup>[14]</sup> Alle basieren allerdings auf Startmaterialien von ungefähr 15 US\$/kg, bestehen aus mindestens 8 Reaktionsschritten und führen zu einer Ausbeute von unter 10 %. Eine Alternative stellt die patentierte Semisynthese dar, bei der die biosynthetisierte Zwischenstufe Amorpha-4,11-dien chemisch zu Artemisinin umgesetzt wird.<sup>[15]</sup> Amorpha-4,11-dien ist das Produkt des ersten Artemisininbiosyntheseschritts (Abbildung I.4). Katalysiert wird er durch die Amorphadiensynthase (ADS), die Farnesyldiphosphat zu Amorpha-4,11-dien zyklisiert.<sup>[16] [17]</sup>

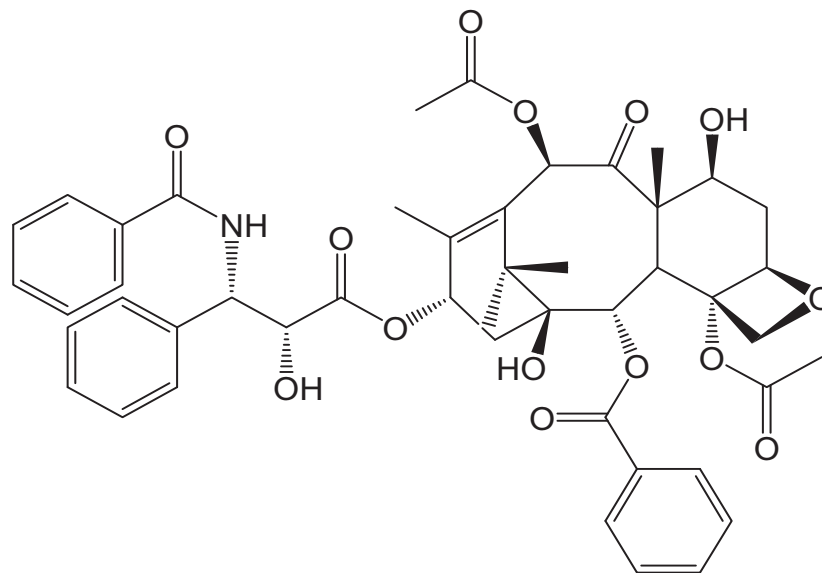


**Abbildung I.4: Schematische Darstellung der Artemisininbiosynthese.** Biosynthese von Amorpha-4,11-dien mit anschließender chemischen Umsetzung zu Artemisinin. **FPP:** Farnesyldiphosphat; **ADS:** Amorphadiensynthase

### I.1.3 Taxol

Taxol ist ein Diterpen, welches heutzutage aufgrund seiner antiproliferativen Wirkung gegen diverse Tumorarten das wohl meist eingesetzte natürliche Krebsmedikament darstellt. Seine Struktur ist in

Abbildung I.5 dargestellt.



**Abbildung I.5: Struktur von Taxol (Paclitaxel).** Chemische Bezeichnung: 5 $\beta$ ,20-Epoxy-1,2 $\alpha$ ,4,7 $\beta$ ,10 $\beta$ ,13 $\alpha$ -hexahydrotaxa-11-2-en-9-on 4,10-diacetat 2-benzoat 13-ester mit (2*R*,3*S*)-*N*-benzoyl-3-phenylisoserin

#### I.1.3.1 Entdeckung und Wirkungsweise

Nachdem die Rate von Infektionskrankheiten durch die Entdeckung von Antibiotika drastisch abgenommen hatte, wurden Herzkrankheiten und Krebs zur häufigsten Todesursache der westlichen Welt und damit zum neuen Forschungsschwerpunkt.<sup>[18]</sup> Durch erste Erfolge chemotherapeutischer Substanzen für die Behandlung humaner Krebserkrankungen wuchs das Interesse an weiteren biologisch aktiven Verbindungen, sodass 1955 ein Programm des US-Amerikanischen *National Cancer Institutes* zur großflächigen Suche solcher Agenzien initiiert wurde. Im Rahmen dessen wurden Pflanzen weltweit auf cytotoxische Substanzen untersucht und 1964 schließlich eine Probe von Stamm und Rinde der pazifischen Eibe *Taxus brevifolia* positiv auf Antitumoraktivität *in vivo* getestet. Obwohl die Verbindung bereits zwei Jahre später isoliert werden konnte, wurde ihre Struktur (

Abbildung I.5) erst 1971 publiziert.<sup>[19]</sup> Die Namensgebung „Taxol“ erfolgte somit vor der vollständigen Strukturaufklärung und basiert auf dem Nachweis seiner Hydroxygruppen.<sup>[20]</sup> Neben

*T. brevifolia* konnte Taxol auch in anderen *Taxus*-Spezies wie *Taxus baccata* und *Taxus cuspidata* nachgewiesen werden.<sup>[21]</sup> Allerdings stellt Taxol nur eins von hunderten beschriebener Taxane dar und macht innerhalb einer Pflanze meist nur einen sehr geringen Anteil der gesamten Taxanfraktion aus.<sup>[22]</sup> In manchen *Taxus*-Spezies wurde gar kein Taxol, sondern ausschließlich andere Taxane entdeckt.<sup>[19]</sup> Aber nicht nur aufgrund dessen war die Entdeckung von Taxol ein glücklicher Zufall. Im Gegensatz zu *T. brevifolia*, in dessen Rindenextrakt ursprünglich die cytotoxische Aktivität von Taxol gegen Krebszellen nachgewiesen werden konnte, synthetisieren *T. baccata* und *T. cuspidata* neben Taxol weitaus größere Mengen an Taxin B. Diese ist eine toxische Verbindung, die Taxol durch sein mehrfach oxygeniertes Taxanringsystem allerdings strukturell sehr ähnlich ist. Aufgrund ihrer gleichen Löslichkeit in Ethanol wären die beiden Substanzen bei Extraktion der Rinden von *T. baccata* bzw. *T. cuspidata* anstelle von *T. brevifolia* simultan aufgereinigt worden. Der Extrakt hätte somit durch den Anteil an Taxol zwar *in vitro* cytotoxische Aktivität gegen Tumorzellen gezeigt, bedingt durch die großen Mengen des gleichzeitig isolierten Herzgifts aber ebenfalls starke *in vivo* Toxizität.<sup>[23]</sup> Das Projekt Taxol wäre demnach aufgrund der tödlichen Nebenwirkungen sehr wahrscheinlich eingestellt worden.

Die wenig später *in vitro* aufgeklärte Wirkungsweise von Taxol machte es trotz der geringen Ausbeute interessant, da sie sich von derjenigen zu der Zeit bekannter pflanzlicher Cytostatika wie Vinblastin, Colchicin oder Podophyllotoxin merklich unterschied.<sup>[24]</sup> Zwar führt Taxol ebenfalls zu einer Störung des Mitosespindelapparats, inhibiert aber nicht die Anlagerung von Tubulineinheiten an die Mikrotubuli, sondern stabilisiert diese durch eine verhinderte Depolymerisation. Die Funktion der Mikrotubuli, die auf einem dynamischen Gleichgewicht zwischen Verlängerung und Verkürzung beruht, wird dadurch gestört und es entstehen abnorme Mikrotubulusstrukturen. Der genaue Prozess, der schlussendlich zur Apoptose der Zelle führt, wurde bisher nicht aufgeklärt. Fest steht jedoch, dass es durch den funktionsunfähigen Spindelapparat während der Mitose zu Chromosomenbrüchen und einer abgebrochenen Zellteilung kommt. Die mit Taxol behandelte Zelle blockiert in der G<sub>2</sub>/M-Phase oder fällt in die G<sub>1</sub>-Phase zurück, um sich erneut zu teilen. Das Ergebnis sind vielkernige Zellen oder Zellen mit zu kleinen Zellkernen, wodurch die Signalkaskade für den mitochondrialen Apoptoseweg ausgelöst werden kann.<sup>[25]</sup> Im Laufe der Zeit wurden weitere Naturstoffe wie Epothilon A und B aus *Sorangium cellulosum*, Eleutherobin aus Weichkorallen der Gattung *Eleutherobia*, sowie Sarcodictyin A und B aus *Sarcodictyon roseum* identifiziert, die einen zu Taxol analogen Wirkungsmechanismus zeigen, in ihrer Wirksamkeit allerdings unterliegen.<sup>[26]</sup>

Die klinischen Phasen wurden von Taxol in den 1980ern zwar erfolgreich durchlaufen, verschlangen aber Hunderttausende von kg an Eibenrinde, sodass die Sorge um ausreichenden

Nachschub an den langsam wachsenden Eibenbäumen aufkam. Das *National Cancer Institute* machte daher 1991 eine Ausschreibung, im Zuge dessen sich die Firma Bristol-Meyers Squibb gegen mehrere andere an Taxol interessierte pharmazeutische Unternehmen durchsetzte und ein *Cooperative Research and development agreement* unterzeichnete. Darin verpflichtete sie sich, die Produktion von Taxol, seine Vermarktung, sowie die Entwicklung alternativer Quellen voran zu treiben. Heute vertreibt Bristol-Meyers Squibb Taxol unter dem generischen Namen Paclitaxel zur Behandlung einer Reihe von Krebsarten wie Eierstock-, Brust-, Prostata- und Lungenkrebs. Mit 22 % Weltmarktanteil ist Taxol heutzutage das meist eingesetzte Krebsmedikament.<sup>[27]</sup> Neben seiner ursprünglichen Verwendung wird Paclitaxel auch bei Verengungen von Herzkranzgefäßen eingesetzt. So konnte gezeigt werden, dass die unerwünschte Neubildung einer Endothelschicht nach Implantation einer Gefäßprothese verhindert wird, wenn Paclitaxel (in einer biologisch abbaubaren Beschichtung integriert) dosiert abgegeben wird.<sup>[28]</sup> Diese PES (*Paclitaxel-eluting stents*) sind die einzigen von der FDA anerkannten DEGs (*Drug-eluting stents*) in den Vereinigten Staaten und sind unter dem Namen TAXUS<sup>®</sup> Express<sup>2™</sup> erhältlich.<sup>[29]</sup>

Neben seiner geringen natürlichen Verfügbarkeit bestand das Hauptproblem von Taxol in seiner schlechten Löslichkeit in Wasser. Die entwickelte Formulierung mit Cremophor EL<sup>®</sup> verbesserte diese zwar, löste allerdings gleichzeitig durch die benötigte hohe Dosierung auch allergische Reaktionen aus,<sup>[30]</sup> die nur durch eine Prämedikation mit Histaminen umgangen werden konnte.<sup>[31]</sup> Auf der Suche nach Verbindungen mit gleichem Wirkungsmechanismus, aber besseren pharmakologischen Eigenschaften wurden daher verschiedene *Taxus* spp. auf weitere Taxane untersucht, die ebenfalls an Tubulin binden und ähnliche biologische Aktivitäten aufweisen. Die in größten Mengen isolierbare Substanz (0,1 % aus Nadeln) stellte sich als 10-Deacetylbaccatin III (Abbildung I.6) heraus. Anschließend Versuche, 10-Deacetylbaccatin III an C10 und C13 zu einer cytotoxischen Substanz zu modifizieren, resultierten in Taxotere (Abbildung I.6). Im Vergleich zu Taxol ergab die Analyse dieser Verbindung eine verbesserte Löslichkeit und höhere Aktivität.<sup>[32]</sup> Seit 1995 wird sie unter dem Namen Docetaxel von Sanofi-Aventis vertrieben.

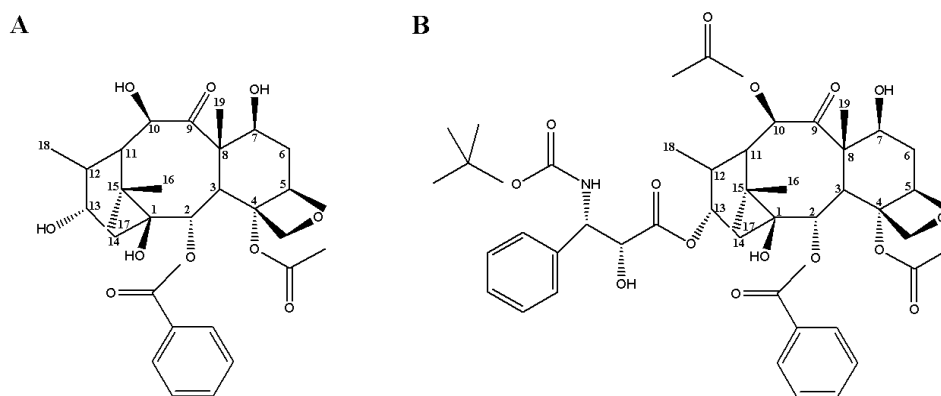


Abbildung I.6: Struktur der extrahierbaren Verbindung 10-Deacetylbaccatin III (A) und dem daraus synthetisierten Taxol-Analogon Taxotere (B)

Für die Entwicklung weiterer Analoga wurde eine Struktur-Aktivitäts-Analyse von Taxol durchgeführt. Mit den identifizierten, für die zytostatische Aktivität essenziellen funktionellen Gruppen sollten anschließend isolierbare Taxane mit besserer Löslichkeit modifiziert werden, um ihnen zusätzlich auch zytostatische Eigenschaften zu vermitteln. Abbildung I.7 fasst die Ergebnisse der Struktur-Aktivitätsstudien zusammen.

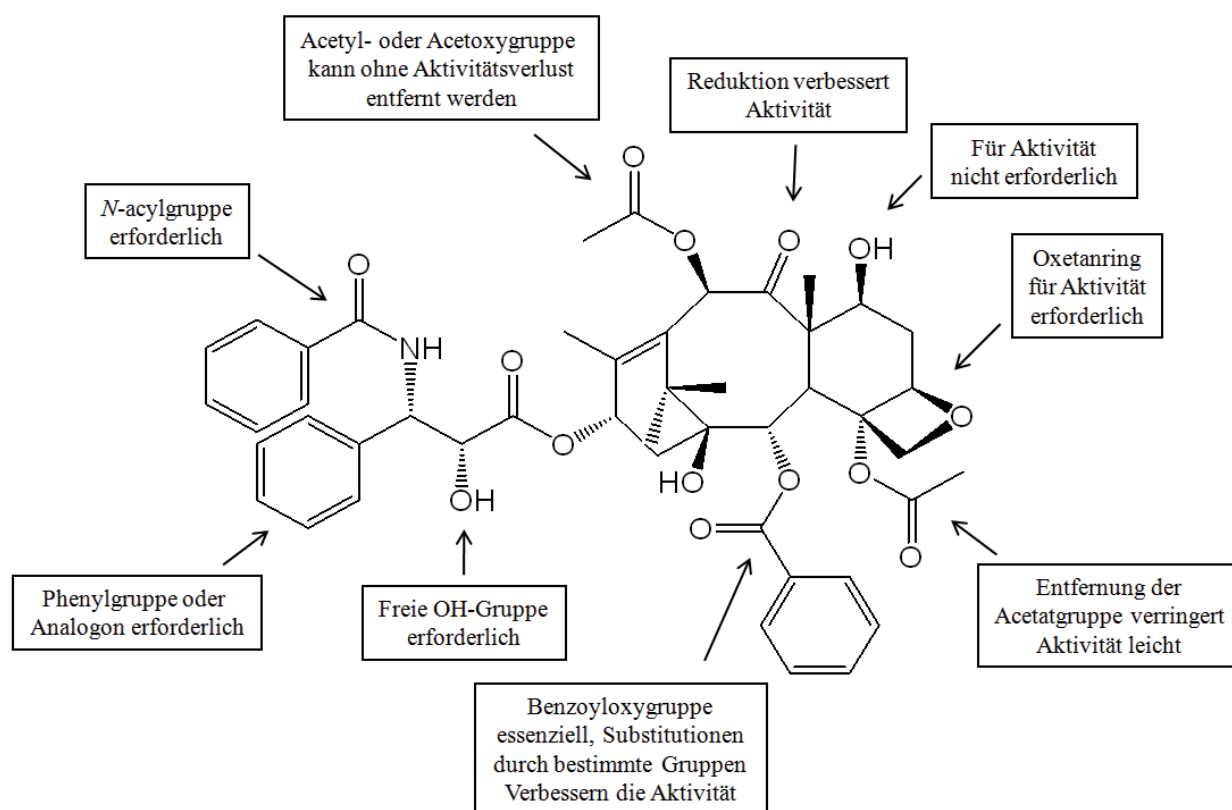


Abbildung I.7: Struktur-Aktivitäts-Beziehungen der Verbindung Taxol<sup>[31]</sup>

Eines von vielen isolierten Taxanen war 14 $\beta$ -Hydroxy-10-deacetylbaccatin III aus *T. wallachiana*, welches durch die zusätzliche Hydroxylierung am C14 eine erhöhte Wasserlöslichkeit zeigte.<sup>[33]</sup> Weitere Modifizierungen führten zu einer Reihe aktiver Verbindungen, von denen eine [13-(*N*-Boc- $\beta$ -isobutylisoserinyl)-14-hydroxybaccatin-1,14-carbonat] aufgrund ihrer 20-fach erhöhten antiproliferativen Aktivität im Vergleich zu Taxol, sowie einer fehlenden Kreuzresistenz von der italienischen Firma Indena für präklinische Studien ausgewählt wurde.<sup>[34]</sup> Durch die orale Verabreichungsmöglichkeit und ausbleibende Toxizitätsproblemen<sup>[35]</sup> wurden die klinischen Phasen 1 und 2 schnell durchlaufen. Allerdings beträgt der Gehalt der Ausgangssubstanz 14 $\beta$ -Hydroxy-10-deacetylbaccatin III in der Pflanze nur 0,04 %.<sup>[33]</sup> Die Funktionalisierung des in größeren Mengen isolierbaren 10-Deacetylbaccatins III an C14 ist zwar möglich, basiert allerdings auf der  $\beta$ -Hydroxylierung gefolgt von einer Carbonylierung des 1,14-Diols mit einer Ausbeute von nur 30 %, <sup>[36]</sup> was nur den ersten Schritt der ganzen Semisynthese darstellt. Aufgrund der geringen Mengen an isolierbarem 14 $\beta$ -Hydroxy-10-deacetylbaccatin III und der schlechten Ausbeute einer semisynthetischen Umsetzung von 10-Deacetylbaccatins III zum gewünschten Produkt, ist es nicht verwunderlich, dass keinerlei Angaben zu Phase III veröffentlicht wurden und Indena sich heutzutage wieder auf die semisynthetische Produktion von Paclitaxel konzentriert.

#### 1.1.3.2 Vorkommen und Gewinnung

Neben *Taxus brevifolia*, in dessen Rinde Taxol ursprünglich detektiert worden war, wurde die Verbindung selbst, sowie andere Taxane auch, in weiteren *Taxus* spp. nachgewiesen,<sup>[37][189][19][22]</sup> wobei Taxol nur einen kleinen Teil der gesamten Taxanmenge ausmacht. Je nach Geschlecht, Spezies und Alter variierte der Gehalt von 0,06 g bis 8,038 g pro Baum.<sup>[38]</sup> Im Vergleich zeigte die Rinde von *Taxus brevifolia* zwar den höchsten Taxolgehalt, aufgrund der geringen Menge und aufwendigen Extraktion würde die Behandlung eines Patienten allerdings immer noch die Rinde mehrerer Bäume benötigen. Eine 20 Jahre alte Eibe produziert bis zu 30 kg an Blättern und 5 kg an Rinde, woraus 4 g Taxol isolierbar sind. Aufgrund des weltweiten Bedarfs von 800 bis 1000 kg Taxol pro Jahr wird die berechnete Grenze zur Erhaltung der Eibe heutzutage bereits um das fünffache überschritten,<sup>[39]</sup> sodass diese Quelle zur Isolierung den enormen Bedarf langfristig nicht decken kann. Auch in Mikroorganismen wie *Taxomyces andreanae*, isoliert aus der inneren Rinde von *T. brevifolia* oder *Pestalotiopsis* spp. aus der Rinde von *T. wallachiana* konnte Taxol nachgewiesen werden.<sup>[40-42]</sup> Allerdings führt der Taxolgehalt von 24-64 ng/L zu keiner kommerziellen Anwendungsmöglichkeit. Die erste Totalsynthese der komplexen Verbindung Taxol wurde bereits 1994 gezeigt,<sup>[43-45]</sup> besteht allerdings aus ca. 40 Reaktionsschritten und ist aufgrund der vielen Stereozentren mit einer Maximalausbeute von nur 2% nicht ökonomisch. Mittlerweile

wurden weitere Synthesewege aufgezeigt, die aufgrund der schlechten Ausbeute ebenfalls eher von akademischem Interesse sind.<sup>[46][47][48][49]</sup> 1988 wurde mit der semisynthetischen Umsetzung der isolierbaren Vorstufe 10-Deacetylbaaccatin III aus den Nadeln die erste Alternative für die Produktion von Taxol gezeigt, die nicht auf der Entfernung der Rinde basiert und daher auf Dauer mit keiner Ausrottung des Eibenbaums einhergehen würde.<sup>[50]</sup> Die anschließende Untersuchung unterschiedlicher Spezies auf den extrahierbaren Gehalt an 10-Deacetylbaaccatin ihrer Nadeln ergab Mengen von 0,009 % (w/w) bis zu 0,02 % (w/w) im besten Produzenten *Taxus baccata*,<sup>[51]</sup> der später durch verbesserte Extraktionsmethoden auf 0,1 % optimiert werden konnte.<sup>[32]</sup> Die Konversion der Vorstufen zu Taxol besteht aus elf synthetischen Schritten, sieben Isolierungen und benötigt 13 Lösungsmittel, sowie 13 organische Reagenzien (Ritter, 2004). Indena, einer der führenden Hersteller, produziert Paclitaxel seit 2007 anhand ihrer patentierten Methode der Semisynthese. Das dafür benötigte 10-Deacetylbaaccatin III wird aus *T. baccata* –Pflanzen ihrer firmeneigenen Plantage isoliert (Pressemitteilung Indena, 29. Mai 2007). Bristol-Meyers-Squibb wechselte dagegen 2002 zu der alternativen Produktionsstrategie der Pflanzenzellkultur. Im Gegensatz zur Extraktion von Nadeln gewährleistet die *large-scale* Fermentation einer Pflanzenzellkultur nicht nur eine von der Jahreszeit unabhängige, schnellere und gleichbleibende Produktionsleistung, sondern reduziert außerdem den energetischen Aufwand und die Anzahl benötigter Lösungsmittel bei der Aufreinigung.<sup>[52]</sup> In Zellkulturen verschiedener *Taxus* spp. wie *T. brevifolia*,<sup>[53]</sup> *T. cuspidata*, *T. canadensis*<sup>[54]</sup> und *T. baccata*<sup>[55][56]</sup> konnte eine Produktion von Taxol nachgewiesen werden, die durch die Zugabe des Pflanzenhormons Methyljasmonat als Induktor<sup>[57]</sup> Werte bis zu 23 mg/l/Tag an Taxanen mit einem Taxolanteil von 13-20 % erreichte.<sup>[58]</sup> Das bei Bristol-Meyers-Squibb verwendete System basiert hingegen auf einer von Phyton Inc. (Ithaca, NY, USA) entwickelten Technologie, die mit Kalli von Nadelzellen der chinesischen Eibe *Taxus chinensis* startet.<sup>[52]</sup> Mit einer Kapazität von 75000 Litern verkörpert sie die weltweit größte Anlage einer Pflanzenzellkultur.<sup>[59]</sup> Die optimierte Semisynthese ausgehend von Baaccatin III oder 10-Deacetylbaaccatin III<sup>[60]</sup> stellt momentan die Hauptquelle für Taxol und das Analogon Taxotere dar.

## **I.2 Biosynthese von Terpenen**

### ***I.2.1 Biosynthese der Vorstufen***

Die Grundgerüste von Terpenen bestehen aus Isoprenoiden, die durch Kondensation der Vorstufen Isopentenylidiphosphat (IPP) und Dimethylallyldiphosphat (DMAPP) synthetisiert werden. Im



Gegensatz zu Terpenen kommen Isoprenoide nicht nur in Pflanzen, sondern ubiquitär in Eubakterien, Archaeobakterien und Eukaryoten vor. Trotz ihrer weiten Verbreitung werden sie alle über zwei unabhängige biosynthetische Routen (Mevalonatweg und DXP-Weg) aus unterschiedlichen Ausgangskomponenten synthetisiert. Die jeweilige Genese ist dabei abhängig von Organismus und Ort der Synthese innerhalb der Zelle.

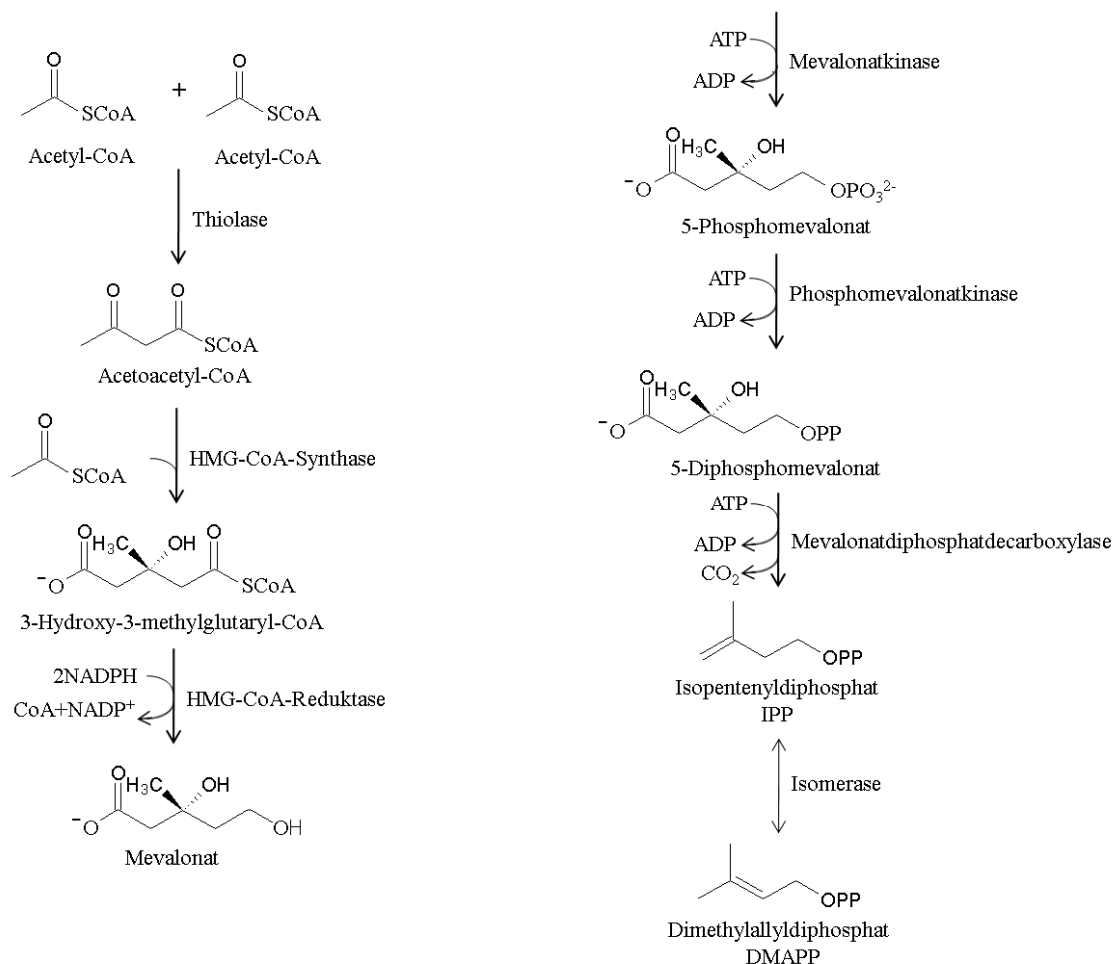
#### *I.2.1.1 Mevalonatweg*

Die Intermediate des Mevalonatwegs wurden bereits in den 50er Jahren anhand von Hefe- und Leberzellextrakten beschrieben.<sup>[61]</sup> Startpunkt ist die Thiolase-katalysierte Kondensation zweier Acetyl-CoA-Moleküle zu Acetoacetyl-CoA. Dieses wird durch Addition einer weiteren Acetyl-CoA-Einheit durch die HMG-CoA-Synthase zu 3-Hydroxy-3-methylglutaryl-CoA (HMG-CoA) umgesetzt (Abbildung I.8). Die anschließende Reduktion der Thioestergruppe von HMG-CoA zum Alkohol wird von der HMG-CoA-Reduktase katalysiert und führt zu Mevalonat. Dieses wird mithilfe von ATP über 5-Phosphomevalonat zu 5-Diphosphomevalonat phosphoryliert und schließlich unter Abspaltung von CO<sub>2</sub> zu IPP decarboxyliert. Eine Isomerase katalysiert die Umwandlung zu DMAPP. Die terminale Doppelbindung macht IPP zu einem schwachen Nucleophil, während DMAPP eher elektrophiler Natur ist. Diese unterschiedlichen Reaktivitäten sind die Basis der Terpenbiosynthese, die mit der enzymkatalysierten Kondensation der beiden Moleküle startet.

Obwohl auch eine *feedback*-Inhibierung der Mevalonatkinase durch GPP und FPP gezeigt werden konnte,<sup>[62]</sup> identifizierten Studien die Umsetzung von HMG-CoA zu Mevalonat durch die HMG-CoA-Reduktase als den geschwindigkeitsregulierenden Schritt des Mevalonatwegs.<sup>[63]</sup> In der Hefe *Saccharomyces cerevisiae* wurden zwei Gene identifiziert (*HMG1* und *HMG2*), die jeweils für ein Isozym der HMG-CoA-Reduktase (*HMG1* und *HMG2*) kodieren.<sup>[64]</sup> Die Enzyme bestehen aus einer identischen cytosolischen katalytischen Domäne, die über eine Linkerregion mit einer im ER verankerten regulatorischen Domäne verbunden ist.<sup>[65]</sup> Diese membrandurchspannenden Regionen der Isozyme zeigen eine Identität von nur 50 %. Der Unterschied der beiden regulatorischen Domänen macht eine sogenannte Kontraregulierung möglich, die bei gleichem Stimulus in unterschiedlicher Aktivität der identischen katalytischen Domäne resultiert.<sup>[66]</sup> <sup>[67]</sup> *HMG1* produziert unter aeroben Bedingungen 83 % der gesamten HMG-CoA-Reduktase-Menge,<sup>[68]</sup> sodass dieses Isozym für Deregulationsstudien ausgewählt wurde. Im Rahmen dieser Studie wurde eine Variante des Enzyms generiert (tHMGR), die *N*-terminal um die membranassoziierte regulatorische Domäne verkürzt und daher löslich ist. Das Fehlen des für die Regulation notwendigen Bereichs resultierte



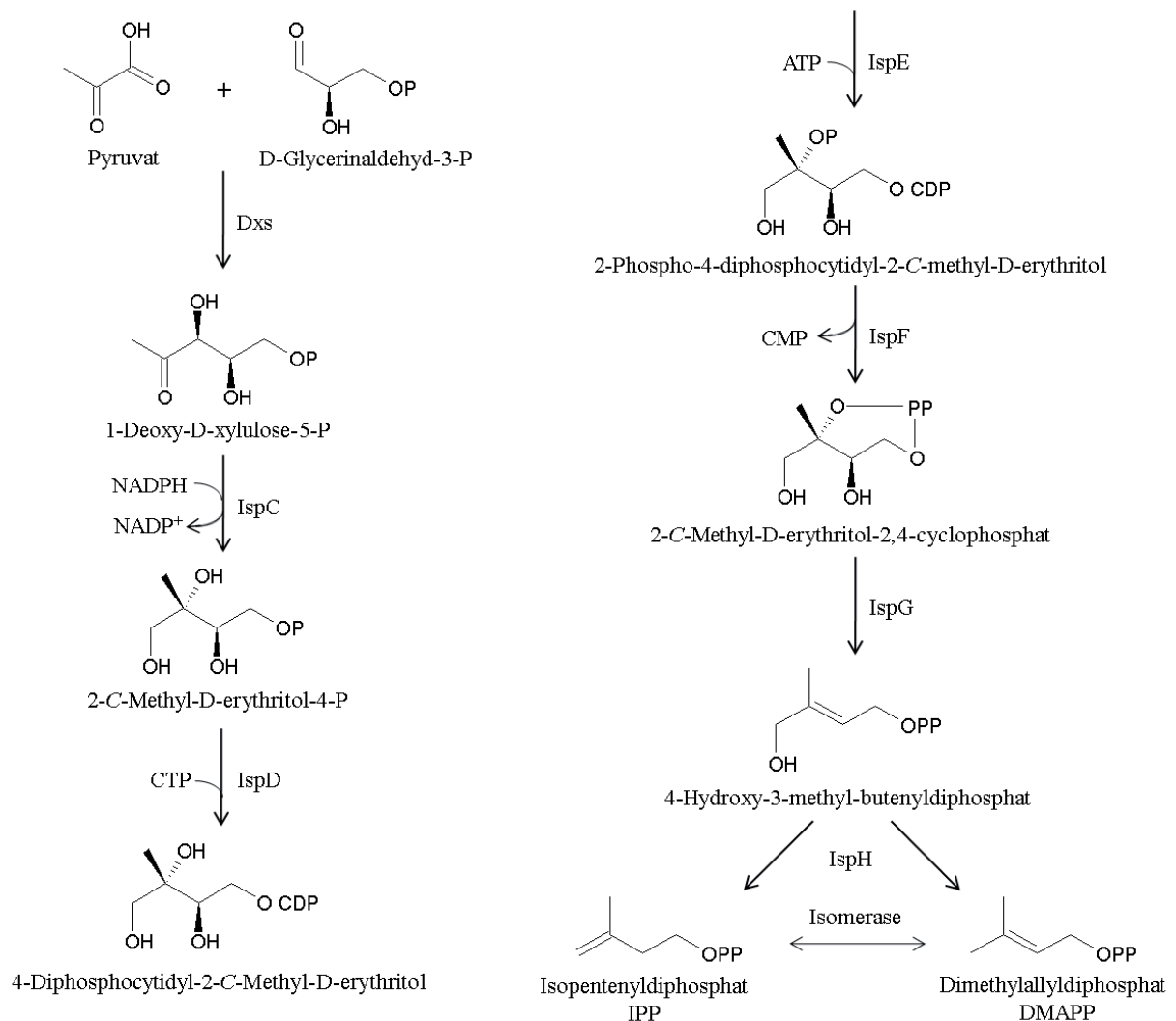
in signifikant höheren Reduktase-mRNS-Mengen, einer 20-fachen Steigerung der Reduktaseaktivität und 8-facher Steroidproduktion.<sup>[69]</sup> Die Regulation der HMG-CoA-Reduktase erfolgt sowohl auf transkriptionaler und translationaler Ebene,<sup>[70]</sup> als auch durch Degradation des Proteins nach Ubiquitinmarkierung.<sup>[71][72]</sup> Des Weiteren wurde eine *Feedback*-Inhibierung durch Produkte des späteren Biosynthesewegs nachgewiesen. Durch Phosphorylierung der HMG-CoA-Reduktase mittels einer AMP-abhängigen Kinase wird die Isoprenoidproduktion zusätzlich an den zellulären Energiehaushalt gekoppelt.<sup>[67]</sup> Grundsätzlich führen alle Regulationsmechanismen bei hohem Bedarf an entsprechenden Molekülen nachfolgender Biosynthesewege zu einer größeren Menge an gebildetem Enzym, bzw. zu einer Verringerung der HMG-CoA-Menge bei sinkender Nachfrage.<sup>[73]</sup>



**Abbildung I.8: Mevalonatweg zur Biogenese der Isoprenvorstufen IPP und DMAPP.** Die Kondensation zweier Acetyl-CoA-Einheiten zu Acetoacetyl-CoA und anschließender Addition einer weiteren AcetylCoA-Einheit führt zu HMG-CoA, welches zu Mevalonat reduziert wird. Nach Phosphorylierung zu Diphosphomevalonat entsteht durch eine Decarboxylierungsreaktion Isopentenylpyrophosphat, welches zu Dimethylallylpyrophosphat isomerisiert wird.

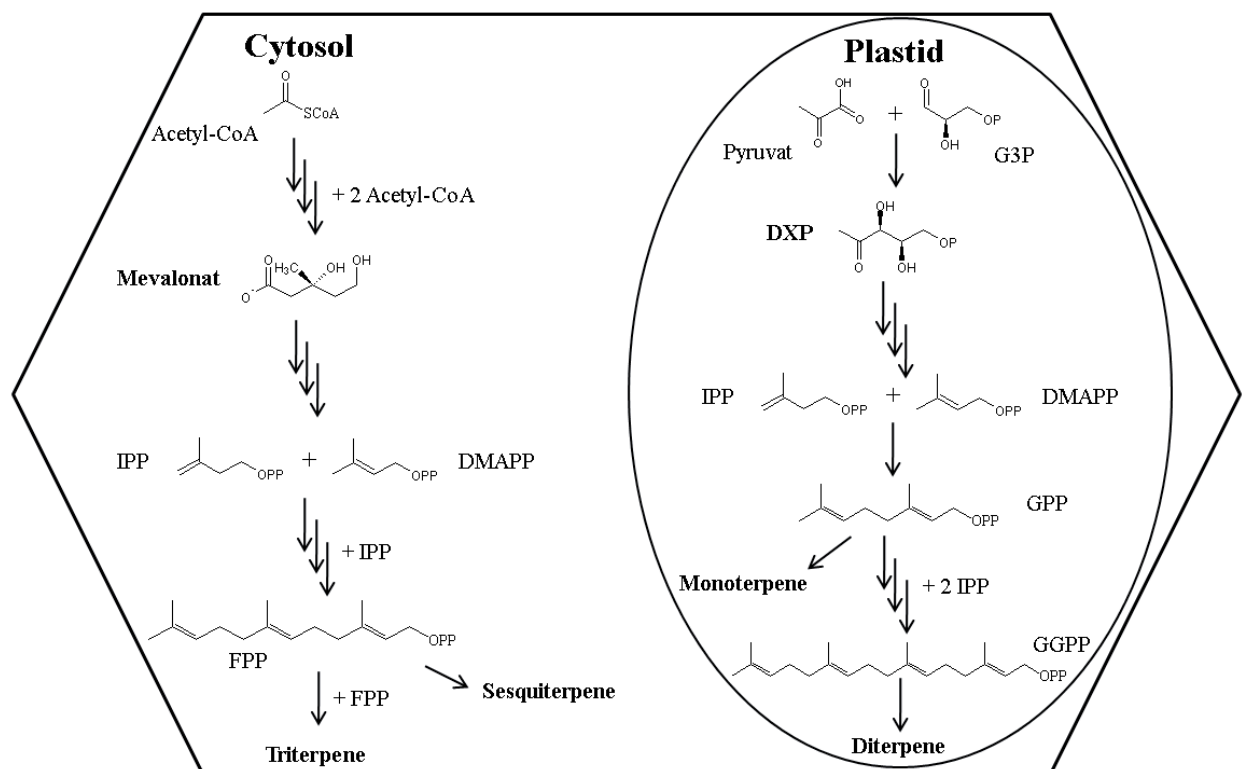
### I.2.1.2 DXP-Weg

Der Mevalonatweg wurde lange Zeit für ubiquitär und einzige Quelle für Terpenvorstufen gehalten,<sup>[74]</sup> da er in zahlreichen Eukaryoten sowie Archaeobakterien nachgewiesen werden konnte. Untersuchungen an Bakterien zur Identifizierung des Mevalonatwegs waren allerdings nicht mit der bekannten Isoprenbiosyntheseroute vereinbar. Versuche zum Einbau <sup>13</sup>C- und <sup>14</sup>C-markierten Acetats in terpenoide Strukturen von Eubakterien wichen vom erwarteten Markierungsmuster für die Synthese dieser Verbindungen über den Mevalonatweg ab. So wurde zum Beispiel <sup>13</sup>C-Acetat nicht in die Seitenkette des Ubichinons von *Escherichia coli* eingebaut.<sup>[75]</sup> Auch konnte keine Wachstumshemmung durch die Zugabe von Mevilonin nachgewiesen werden, welches die HMG-CoA-Reduktase spezifisch hemmt, die den regulierenden Schritt des Mevalonatwegs katalysiert.<sup>[75]</sup> Ebenso wurde bei *Agrobacterium tumefaciens*, *Azotobacterium vinelandii* und *Pseudomonas* spp. kein Einbau von <sup>14</sup>C-Mevalonat, -Formiat oder -Acetat in Ubichinon beobachtet.<sup>[76]</sup> Um die Herkunft bakterieller Terpenoide zu bestimmen, untersuchten Rohmer und Mitarbeiter daher acht verschiedene prokaryotische Organismen auf den Einbau <sup>13</sup>C-markierten Acetats und Pyruvats, sowie Glucose und Erythrose in die jeweiligen Tri- und Polyterpenoide.<sup>[77]</sup> Dazu wurden gezielt Bakterien mit verschiedenen Metabolisierungsrouten der jeweilig zugefütterten Kohlenstoffquelle ausgewählt. Das Ergebnis war der Hinweis auf einen alternativen Stoffwechselweg zur Biosynthese von Isopreneinheiten, der von Pyruvat und einem Triosephosphat ausgeht, welches ein paar Jahre später als Glycerinaldehyd-3-phosphat (GAP) identifiziert werden konnte. Die Kondensation von GAP mit Pyruvat führt zum ersten und namensgebenden Zwischenprodukt dieses Biosynthesewegs 1-deoxy-D-xylulose-5-phosphat (DXP).<sup>[78]</sup> Die Reaktion wird von der DXP-Synthase (Dxs) katalysiert. Es folgt eine intramolekulare Umlagerung mit anschließender Reduktion zu 2C-Methyl-D-erythritol-4-phosphat (MEP) durch die 1-Deoxy-D-xylulose-5-phosphatreduktoisomerase IspC (Abbildung I.9). Die von IspD (4-Diphosphocytidyl-2-C-methyl-D-erythritolsynthase) katalysierte Reaktion von MEP mit Cytidintriosephosphat (CTP) führt zum 4-Diphosphocytidyl-2-C-methyl-D-erythritol. Dieses wird in einer ATP-abhängigen Reaktion durch IspE zu 2-Phospho-4-diphosphocytidyl-2-C-methyl-D-erythritol phosphoryliert und unter Verlust von Cytidinphosphat durch IspF ins zyklische Phosphoanhydrid 2-C-Methyl-D-erythritol-2,4-zyklophosphat überführt. Die anschließende Reduktion zu 4-Hydroxy-3-methyl-butenyldiphosphat wird durch IspG, die Umwandlung zu IPP und DMAPP durch IspH katalysiert. Obwohl beide Isomere koproduziert werden, besteht wie beim Mevalonatweg die Möglichkeit der Umwandlung ins andere Isomer durch eine Isomerase. Das gewährleistet ein ausgeglichenes Verhältnis von IPP und DMAPP, welche bei der Synthese der jeweiligen Terpenvorstufen Farnesyldiphosphat, Geranylgeranyldiphosphat, etc. in unterschiedlicher Stöchiometrie benötigt werden.



**Abbildung I.9: DXP-Weg zur Biosynthese der Isopreneinheiten IPP und DMAPP.** Das Enzym DXP-Synthase katalysiert mit der Kondensation von Pyruvat und Glycerinaldehyd-3-phosphat zu DXP den ersten Schritt des Stoffwechselwegs. DXP wird anschließend zu MEP umgewandelt, welches nach Reaktion mit CTP und Phosphorylierung mittels ATP durch IspE zu 2-Phospho-4-diphosphocytidyl-2-C-methyl-D-erythritol führt. IspF katalysiert die Zyklisierungsreaktion zum Phosphoanhydrid 2-C-Methyl-D-erythritol-2,4-cyclophosphat unter Verlust von CMP. Mittels Reduktion entsteht 4-Hydroxy-3-methyl-butenyldiphosphat, welches durch IspH vornehmlich in Isopentenyl-, aber auch in sein Isomer Dimethylallyldiphosphat umgewandelt wird. **DXP:** 1-Deoxy-D-xylulose-5-phosphat; **MEP:** 2-C-Methyl-D-erythritol-4-phosphat; **Dxs:** 1-Deoxy-D-xylulose-5-phosphatsynthase; **IspC:** 1-Deoxy-D-xylulose-5-phosphatreduktisomerase; **IspD:** 4-Diphosphocytidyl-2-C-methyl-D-erythritolsynthase; **IspE:** 4-Diphosphocytidyl-2-C-methyl-D-erythritolkinase; **IspF:** 2-C-Methyl-D-erythritol-2,4-cyclodiphosphatsynthase; **IspG:** 4-Hydroxy-3-methylbutenyl-diphosphatsynthase; **IspH:** 4-Hydroxy-3-methylbutenyl-diphosphatreduktase.

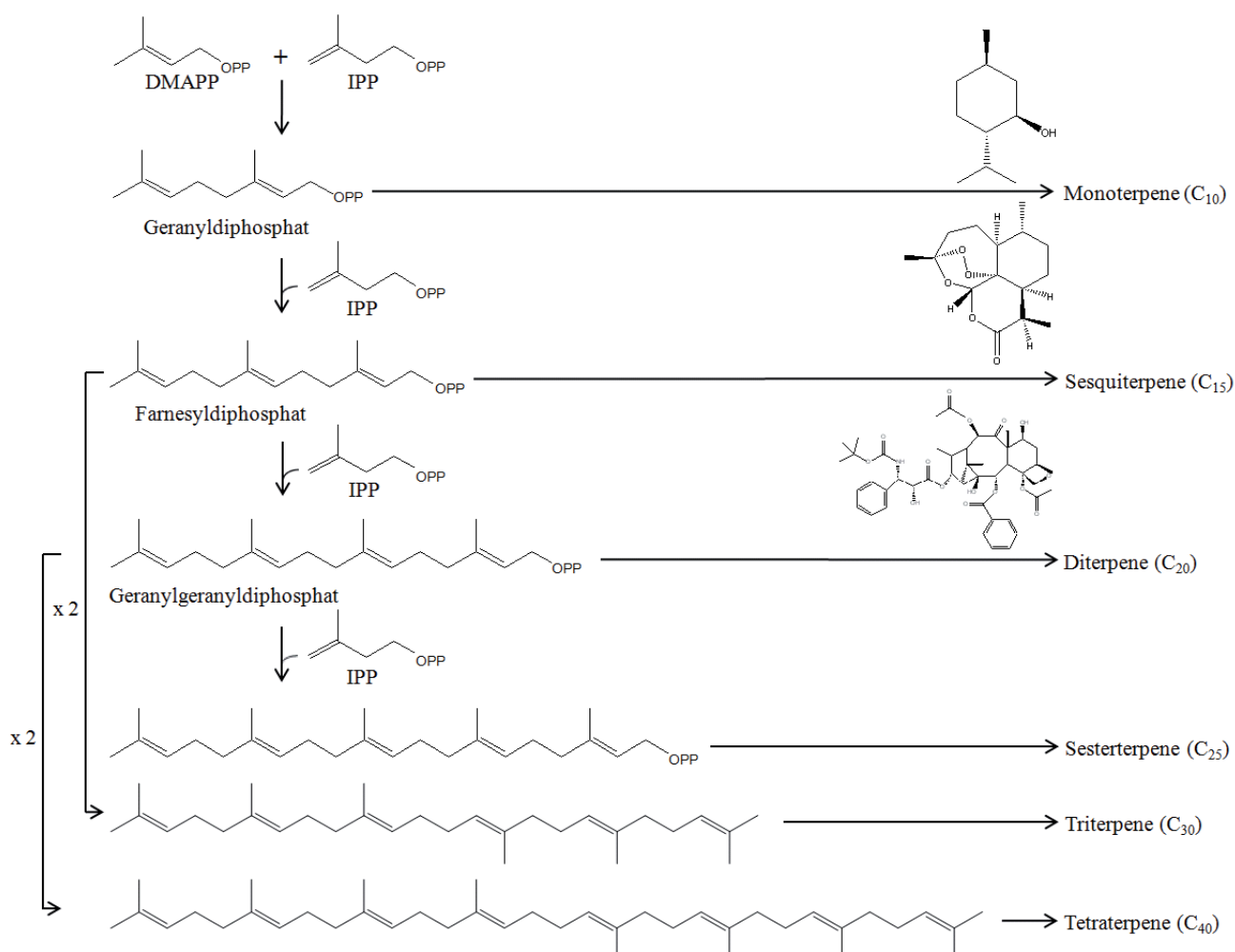
Mittlerweile ist bekannt, dass Archaeobakterien, Pilze und Tiere ausschließlich den Mevalonatweg verwenden.<sup>[79]</sup> In der Mehrheit aller Eubakterien, Cyanobakterien, sowie in bestimmten Algen wurde hingegen der DXP-Weg nachgewiesen. In höheren Pflanzen wurden beide Biosynthesewege identifiziert (Abbildung I.10). Über den cytosolischen Mevalonatweg werden dort Sesquiterpene und Steroide synthetisiert, während der plastidäre DXP-Weg zur Biosynthese von Mono-, Diterpenen, Carotinoiden, Phytol und Plastochinon verwendet wird.<sup>[80]</sup> So erfolgt auch die Biosynthese des Taxangerüsts aus Isoprenoidvorstufen des DXP-Wegs.<sup>[81]</sup>



**Abbildung I.10: Schematische Darstellung der Verteilung von Mevalonatweg (Cytosol) und DXP-Weg (Plastiden) in der Pflanzenzelle.** IPP: Isopentenylidiphosphat; DMAPP: Dimethylallyldiphosphat; FPP: Farnesylidiphosphat; G3P: Glycerinaldehyd-3-phosphat; DXP: 1-Deoxy-D-xylulose-5-phosphat; GPP: Geranylidiphosphat; GGPP: Geranylgeranylidiphosphat

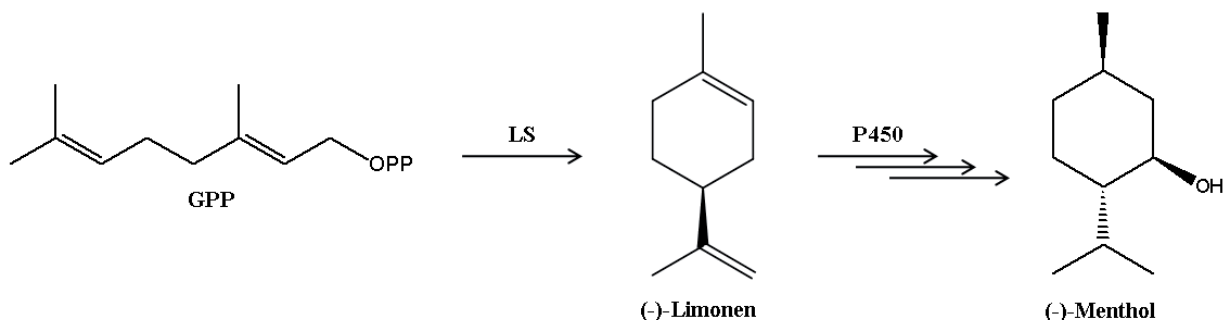
### I.2.2 Biosynthese der Terpengrundgerüste

Die jeweiligen Vorstufen der Terpene bestehen aus den Isoprenoidvorstufen Isopentenylidiphosphat (IPP) und seinem Isomer Dimethylallyldiphosphat (DMAPP). Diese werden über einen nukleophilen Angriff der IPP-Methylengruppe auf die Allyl-CH<sub>2</sub>-Gruppe des DMAPP zu Geranyldiphosphat, der Ausgangssubstanz für alle Monoterpene, verknüpft. Die Weiterreaktion mit zusätzlichen IPP-Einheiten liefert Farnesyldiphosphat, die Vorstufe der Sesquiterpene, und Geranylgeranyldiphosphat, wovon sich alle Diterpene ableiten. Die Verknüpfung zweier Äquivalente Farnesyldiphosphats führt zu Triterpenen, die zweier Geranylgeranyldiphosphatmoleküle zu Polyprenyldiphosphat und somit Tetra- und Polyterpenen (Abbildung I.11).



**Abbildung I.11: Terpenoidbiosynthese.** Mittels Kopf-Schwanz-Kondensation von DMAPP und IPP wird Geranyldiphosphat gebildet, die Vorstufe aller Monoterpene (C<sub>10</sub>). Durch die Addition weiterer IPP-Moleküle entsteht Farnesyldiphosphat (Sesquiterpen, C<sub>15</sub>), Geranylgeranyldiphosphat (Diterpene, C<sub>20</sub>), sowie die Vorstufe der Sesterterpene (C<sub>25</sub>). Durch Kondensation zweier Farnesyldiphosphats entsteht das Grundgerüst der Triterpene (C<sub>30</sub>), durch die Kondensation zweier Geranylgeranyldiphosphats dasjenige von Tetraterpenen (C<sub>40</sub>).

In den meisten Fällen werden die synthetisierten Diphosphat-Vorstufen (Abbildung I.11) anschließend durch eine Terpensynthese zyklisiert und das entstandene Terpengrundgerüst weiter funktionalisiert. Diese Modifikationen sind hauptsächlich Cytochrom-P450-Monooxygenase-katalysierte Reaktionen.<sup>[6]</sup> Abbildung I.12 gibt eine schematische Darstellung am Beispiel Menthol wieder.



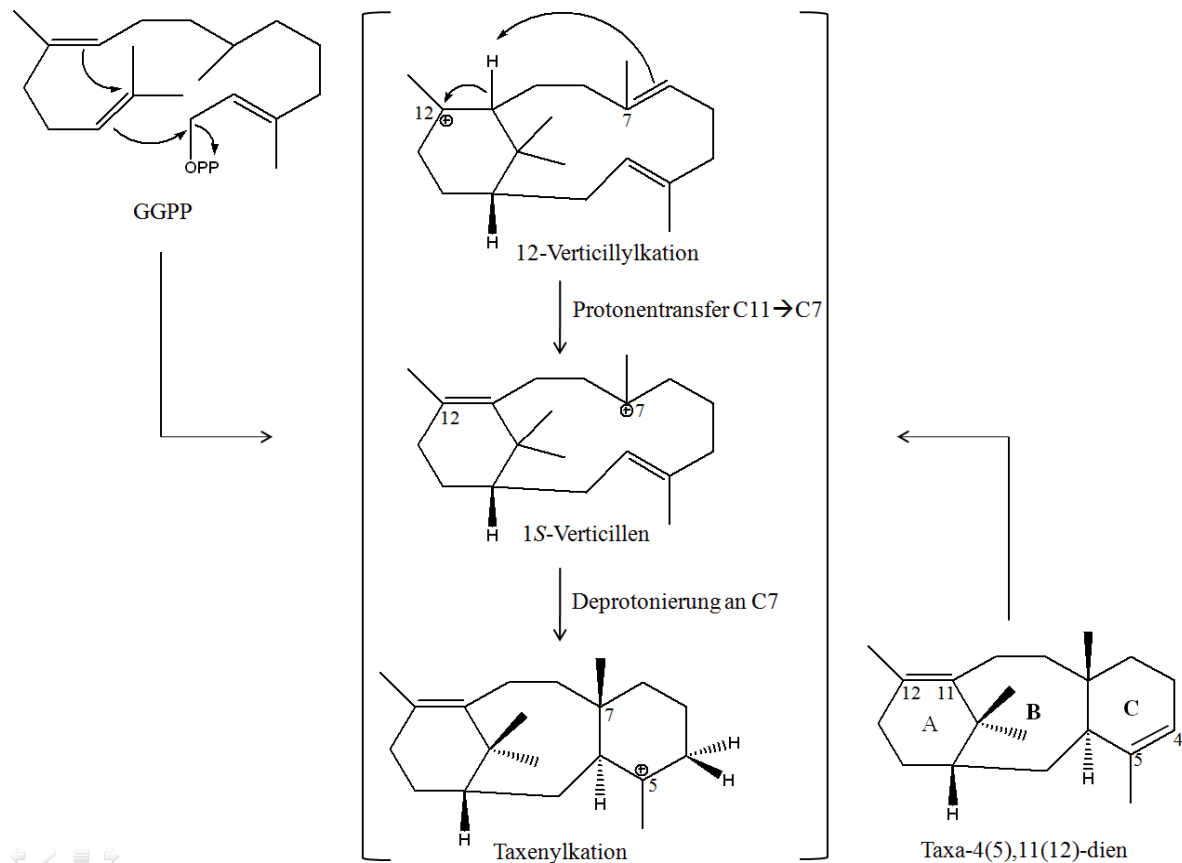
**Abbildung I.12: Schematische Darstellung der Biosynthese von Terpenen am Beispiel von Menthol.** Die durch Kondensation der isoprenoiden Vorstufen IPP und DMAPP generierten Diphosphate werden durch eine Terpensynthese zyklisiert und mittels Cytochrom-P450-Monooxygenase-katalysierte Reaktionen modifiziert. **LS:** Limonensynthase; **P450:** Cytochrom-P450-Monooxygenasen

### I.2.3 Biosynthese von Taxol

#### I.2.3.1 Biosynthese des Taxangrundgerüsts

Die aus ca 19 enzymatischen Schritten bestehende Biosynthese von Taxol startet mit der Zyklisierung des Diterpenvorläufers GGPP zu Taxa-4(5),11(12)-dien. Dies wurde anhand von Umsetzungsversuchen radioaktiv markierten GGPPs durch zellfreie *Taxus brevifolia*-Extrakte gezeigt und durch Inkorporationsstudien des aufgereinigten Produkts in Taxol bestätigt.<sup>[82]</sup> Das katalysierende Enzym Taxadiensynthase wurde aus Eibentrieben partiell isoliert<sup>[83]</sup> und für die weitere Aufklärung des Reaktionsmechanismus verwendet.<sup>[84]</sup> Viele natürliche taxoide Verbindungen zeigen Doppelbindungen an der C4-C20-Position,<sup>[85]</sup> sodass die Vermutung nahe lag, auch im Falle des Taxols könnte ein initiales Grundgerüst in Form von Taxa-4(20),11(12)-dien entstehen, welches anschließend zu dem bereits detektierten Zyklisierungsprodukt Taxa-4(5),11(12)-dien umgeformt wird. Allerdings bestätigten neben Lin und Mitarbeitern auch andere Gruppen die Konversion von GGPP durch seine Ionisierung und Zyklisierung zum Verticillylkation.<sup>[84][86][87][88]</sup> Anhand von an der C10-Position markierten GGPPs konnte gezeigt

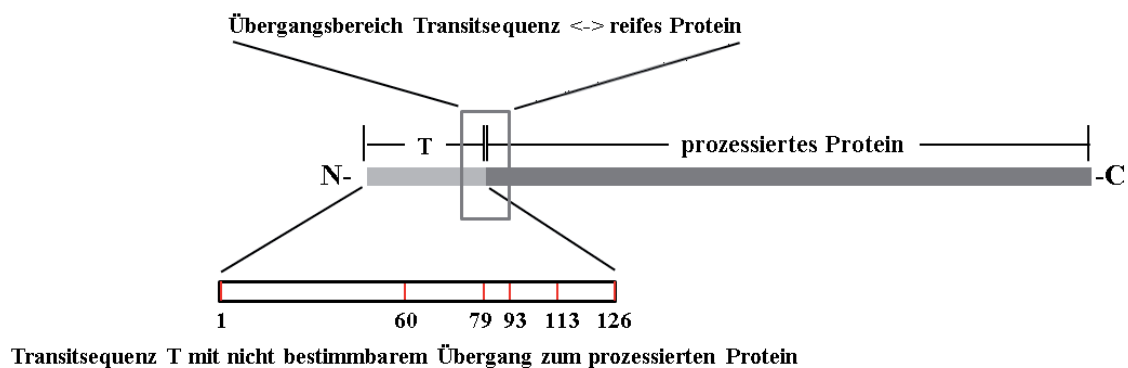
werden, dass der anschließende intramolekulare Protonentransfer von C11 des Verticillylkations (also C10 des GGPP-Moleküls) zu C7 zum Taxenylkation führt.<sup>[84]</sup> Dieses wird am C5 zu Taxa-4(5),11(12)-dien deprotoniert. Taxa-4(20),11(12)-dien als Zwischenstufe der Zyklisierung von GGPP zu Taxa-4(5),11(12)-dien wurde nicht bestätigt.



**Abbildung I.13: Zyklisierungsreaktion von GGPP zu Taxadien.** GGPP zum 12-Verticillylkation zyklisiert, welches durch einen Protonentransfer von C11 zu C7 zum 1S-Verticillin umgewandelt wird. Die Deprotonierung an C7 führt zum Taxenylkation, welches seinerseits an C5 zu Taxa-4(5),11(12)-dien deprotoniert wird (Lin et al, 1996). **GGPP:** Geranylgeranyldiphosphat

Anhand einer cDNS-Bibliothek aus Gesamt-RNS von *T. brevifolia* wurde schließlich die Sequenz des Taxadiensynthasegens entschlüsselt.<sup>[89]</sup> Durch seine heterologe Expression konnte gezeigt werden, dass neben dem Hauptprodukt Taxa-4(5),11(12)-dien (94 %) in geringem Maße auch sein Isomer Taxa-4(20),11(12)-dien (4,8 %) und Vertizillen (1,2 %) entstehen.<sup>[86]</sup> Die erstmals detektierten Nebenprodukte sind durch abnorme Deprotonierungen während der Zyklisierungsreaktion zu erklären. So führt die direkte Deprotonierung des Verticillylkations zum nachgewiesenen Vertizillen, wohingegen das Taxa-4(20),11(12)-dien-Isomer durch die Deprotonierung der C20 Methylgruppe am Taxenylkation entstehen kann.<sup>[86]</sup>

Allerdings wurden nicht nur die Zyklisierungsreaktion und das Produktspektrum des Enzyms analysiert, sondern auch seine Länge nach Prozessierung untersucht. Aus *Taxus brevifolia* isoliert besitzt es eine Größe von 79 kDa.<sup>[83]</sup> Wie auch andere pflanzliche Mono- und Diterpensynthasen wird die Taxadiensynthase als Präprotein mit einer *N*-terminalen Signalsequenz für den Transport in die Plastiden translatiert. Diese wird dort proteolytisch abgespalten und entlässt das reife Protein. Obwohl der genaue Übergang von Signalsequenz zum reifen Protein nicht identifiziert werden konnte, da das prozessierte Protein *N*-terminal blockiert wird, ergab eine Schätzung der heterolog exprimierten entsprechenden cDNS aus *Taxus brevifolia* eine Transitpeptidlänge von 137 Aminosäuren und ein reifes Protein von 862 Aminosäuren. Zusammen zeigten sie eine Größe von 98 kDa.<sup>[89]</sup> Ein Vergleich verschiedener Signalsequenzen bekannter Länge ergab keine konservierten Bereiche, sondern auffällig unterschiedliche Aminosäuresequenzen.<sup>[90]</sup> Daher wurden, um die Länge der prozessierten Taxadiensynthase zu bestimmen, von der Gruppe um Williams verschiedene, über den Bereich des Übergangs variierend verkürzte TDS-Varianten rekombinant in *E. coli* exprimiert und auf ihre Aktivität getestet.<sup>[91]</sup> Untersuchungen des Volllängeproteins, sowie nach Verkürzungen um 60 und 79 Aminosäuren (Abbildung I.14) ergaben detektierbare Taxadiensynthaseaktivität, während Verkürzungen um 93, 113 oder 126 Aminosäuren zu einer inaktiven Form des Enzyms führten.



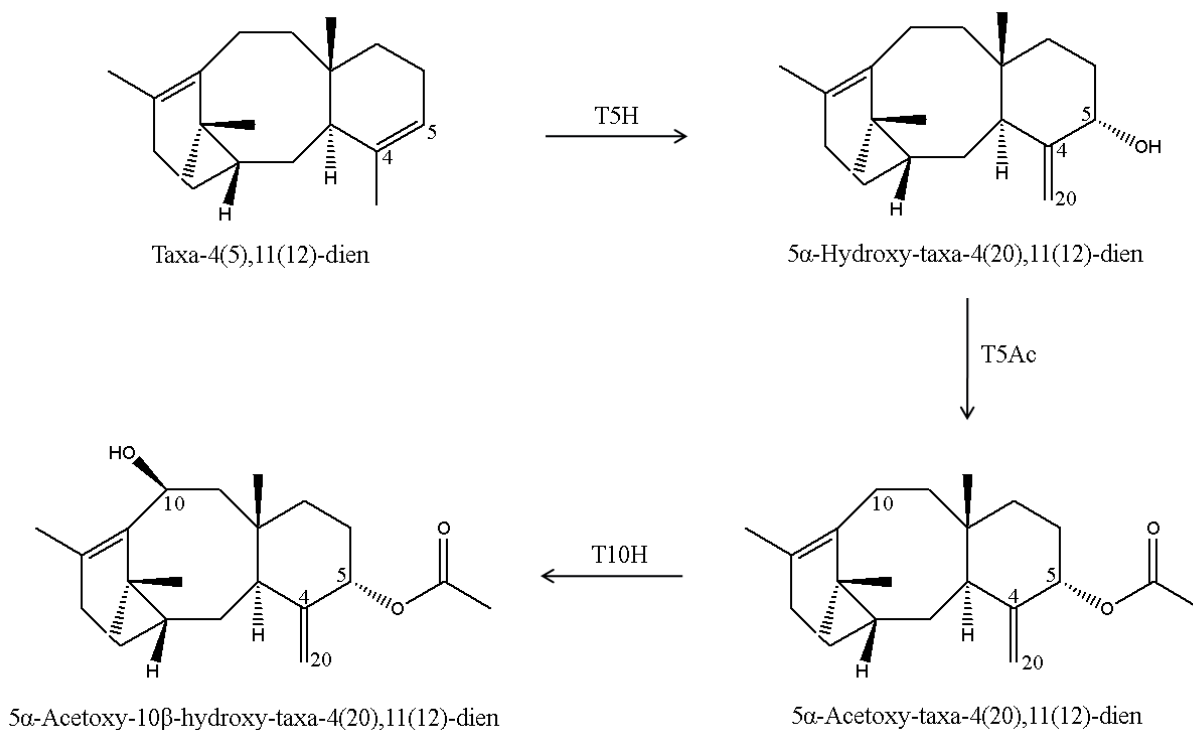
**Abbildung I.14:** Schematische Darstellung der Taxadiensynthase bestehend aus einer *N*-terminalen Transitsequenz und dem prozessierten Protein. **Dunkelgrau:** ungefähre Bereich des reifen Proteins; **Hellgrau:** ungefähre Bereich der Transitsequenz; **Kasten:** Übergangsbereich zwischen Transitsequenz und reifem Protein; **Rot:** Stellen innerhalb der Transitsequenz, an denen die Taxadiensynthase für die rekombinante Expression des reifen Proteins in *E. coli* verkürzt wurde.<sup>[91]</sup> **1-126:** jeweilige *N*-terminale Verkürzung (1=Volllängeprotein)



### I.2.3.2 Modifizierung des Taxangrundgerüsts

Taxol ist ein komplexes Diterpenoid mit diversen Oxygenierungen (Abbildung I.5), die wie auch bei anderen terpenoiden Verbindungen<sup>[92]</sup> von Cytochrom-P450-Monooxygenasen katalysiert werden.<sup>[93][94]</sup> Die Gruppe der P450-Cytochrome sind Hämoproteine, die an der oxidativen Transformation sowohl endogener als auch exogener Moleküle beteiligt sind. Im Zuge dessen inserieren sie ein Atom molekularen Sauerstoffs in das jeweilige Substrat, während das andere Sauerstoffatom zu Wasser reduziert wird. Die benötigten Elektronen stammen von NADPH, welches mit einem Flavoprotein (NADPH:Cytochrom-P450-Reduktase) interagiert, das die Elektronen von NADPH auf das jeweilige Cytochrom-P450-Protein überträgt.<sup>[95]</sup> Auf diese Weise katalysieren sie die verschiedensten Reaktionen von Hydroxylierungen, über Epoxidierungen, Dealkylierungen, Isomerisationen, Dimerisationen, Dehydratationen, C-C-Spaltungen, Decarboxylierungen, Oxidation von Schwefel und Stickstoff, bis hin zu Dehalogenierungen und Desaminierungen.<sup>[96]</sup> Cytochrom-P450-Monooxygenasen wurden sowohl in allen Eukaryoten als auch in zahlreichen Prokaryoten nachgewiesen. Die eukaryotischen Enzyme sind membranständig. Die Verankerung erfolgt dabei durch eine *N*-terminale Transmembranhelix, die das Protein in der ER-Membran fixiert,<sup>[97]</sup> während der Rest des Proteins im Cytosol verbleibt. Daran assoziiert befindet sich die ähnlich verankerte, für den Elektronentransfer von NADPH in das katalytische Zentrum des P450-Proteins benötigte NADPH-Reduktase.<sup>[96]</sup> Um P450-Cytochrome des Taxolbiosynthesewegs zu finden, wurde daher die Mikrosomenfraktion von *T. brevifolia* bzw. *T. cuspidata* verwendet, um sie auf ihre Fähigkeit, radioaktiv markiertes Taxa-4(5),11(12)-dien in ein polareres Produkt umzuwandeln, zu testen.<sup>[98]</sup> Oxygenierte Taxane mit einer Doppelbindung an der C5-Position sind nicht bekannt. Dagegen ist ein großer Anteil von ihnen C5-hydroxyliert. Es wurde daher angenommen, dass eine Hydroxylierung an dieser Position auch hier der nächste Schritt sein könnte. Die entstandene Verbindung wurde daher mit dem synthetischen 5 $\alpha$ - als auch 5 $\beta$ -Epimer des Taxa-4(5),11(12)-dien-5ols verglichen. Ein Abgleich mit Retentionszeit und Massenspektrum der authentischen Verbindungen zeigte, dass Taxa-4(5),11(12)-dien durch die Pflanzenmikrosomen zu Taxa-4(20),11(12)-dien-5 $\alpha$ -ol hydroxyliert wurde. Inkorporationsversuche bestätigten die Verbindung als Vorstufe komplexer Taxane wie 10-Deacetylbaaccatin III, Cephalomanin und Taxol.<sup>[98]</sup> Erstaunlicherweise stellte sich in späteren Untersuchungen das Isomer Taxa-4(20),11(12)-dien als vergleichbar effizientes Substrat sowohl von rekombinant exprimierter 5 $\alpha$ -Hydroxylase, als auch von der Mikrosomenfraktion induzierter *Taxus*-Zellen dar. Das führte zu dem Schluss, dass die Cytochrom-P450-vermittelte Konversion von Taxa-4(5),11(12)-dien zu Taxa-4(20),11(12)-dien-5 $\alpha$ -ol nicht über ein Epoxid, sondern über ein radikales Zwischenprodukt

verläuft, welches am C5 hydroxyliert wird, wobei eine 4,20 Doppelbindung entsteht.<sup>[99]</sup> Die ersten Schritte des Taxolbiosynthesewegs sind in Abbildung I.15 dargestellt.



**Abbildung I.15: Frühe Schritte der Taxolbiosynthese.** **T5H:** Taxadien-5α-hydroxylase<sup>[98]</sup>; **T5Ac:** Taxadien-5α-O-acetyltransferase<sup>[100]</sup>; **T10H:** Taxan-10β-hydroxylase<sup>[101]</sup>

Viele Taxoide zeigen eine Acylgruppe an der C5-Position, sodass eine Acetylierung von Taxa-4(5),11(12)-dien-5α-ol als anschließender Reaktionsschritt der Taxolbiosynthese in Betracht gezogen wurde. Die Bestätigung erfolgte durch Homogenisate von Methyljasmonat-induzierten *T. canadensis*-Zellkulturen. Diese setzten radioaktiv markiertes Taxa-4(5),11(12)-dien-5α-ol in einer Acetyl-CoA-abhängigen Reaktion zu einer Verbindung um, die authentischem Taxa-4(20),11(12)-dien-5α-yl-acetat entsprach. Anhand ihrer Aktivität konnte die Acetyl-CoA:Taxadienol-O-transferase partiell aufgereinigt und charakterisiert werden.<sup>[100]</sup>

Um die zahlreichen weiteren P450 Monooxygenasen der Taxolbiosynthese zu finden und voneinander unterscheiden zu können, wurden cDNS-Bibliotheken induzierter und nicht induzierter *Taxus*-Zellen verglichen. Da bereits gezeigt worden war, dass in die Taxolbiosynthese involvierte Gene durch die Zugabe von Methyljasmonat auf transkriptioneller Ebene aktiviert werden,<sup>[57][102]</sup> konnten mit entsprechenden P450-spezifischen Primern amplifizierte Sequenzen selektiert, in Hefen

exprimiert und anhand ihrer Funktionalität untersucht werden. Auf diesem Wege wurden 13 Volllänge-Cytochrom-P450-Sequenzen gefunden, die mittels Zufütterung verschiedener taxoider Substanzen *in vivo* auf ihre Oxygenaseaktivität hin untersucht wurden.<sup>[101]</sup> Die genaue Abfolge der noch verbleibenden Oxygenierungen des Taxangrundgerüsts im Rahmen der Taxolbiosynthese war nicht bekannt, wurde durch die relative Häufigkeit von Taxoiden mit entsprechenden funktionellen Gruppen jedoch auf die Reihenfolge C10, C2, C9 und schließlich C13 vermutet, während die Oxygenierungen an C7 und C1 im späteren Verlauf eingefügt würden. So erstaunte es auch nicht, dass anhand der cDNS-Bibliothek eine Cytochrom-P450 Oxygenase gefunden wurde, die radioaktiv markiertes Taxa-4(20),11(12)-dien-5 $\alpha$ -yl-acetat an der C10-Position hydroxyliert und in Taxa-4(20),11(12)-dien-5 $\alpha$ -acetoxy-10 $\beta$ -ol umwandelt.<sup>[101]</sup>

Eine alternative Route ausgehend von Taxa-4(20),11(12)-dien-5 $\alpha$ -ol wurde von Wheeler und Mitarbeitern vorgeschlagen. Die Mikrosomenfraktion induzierter *Taxus*-Zellen wurden sowohl mit Taxa-4(20),11(12)-dien-5 $\alpha$ -ol, als auch mit Taxa-4(20),11(12)-dien-5 $\alpha$ -yl-acetat inkubiert und die entsprechenden Produkte charakterisiert.<sup>[103]</sup> Während letzteres wie bereits gezeigt<sup>[101]</sup> zu Taxa-4(20),11(12)-dien-5 $\alpha$ -acetoxy-10 $\beta$ -ol umgewandelt wurde, führte das eingesetzte Taxa-4(20),11(12)-dien-5 $\alpha$ -ol zu einer Konversion in Taxa-4(20),11(12)-dien-5 $\alpha$ , 13 $\alpha$ -diol und damit einer alternativen Route der Taxoidsynthese.<sup>[103]</sup> Die Theorie der Existenz mehrerer Stoffwechselwege zu Taxol wurde wenig später erhärtet, als die identifizierte 13 $\alpha$ -Hydroxylase Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ -diol-5-acetat nicht zu Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ ,13 $\alpha$ -triol-5-acetat umsetzte. Die Inkubation der 10 $\beta$ -Hydroxylase mit Taxa-4(20),11(12)-dien-5 $\alpha$ ,13 $\alpha$ -diol führte ihrerseits ebenfalls zu keiner Umsetzung.<sup>[104]</sup>

Um weitere Reaktionsschritte des Taxolbiosynthesewegs aufzuklären, wurden die beiden bekannten funktionalisiertesten Produkte der alternativen Biosynthesewege Taxa-4(20),11(12)-dien-5 $\alpha$ ,13 $\alpha$ -diol und Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ -diol-5-acetat radioaktiv markiert und die cDNS-Bibliothek<sup>[101]</sup> verwendet, um anhand einer Substratumsetzung die entsprechenden Enzyme zu finden.<sup>[105]</sup> Dabei wurde ein Klon gefunden, dessen Produkt Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ -diol-5-acetat in eine polarere Verbindung konvertierte, die als Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ ,14 $\beta$ -triol-5-acetat identifiziert werden konnte. Im Gegensatz dazu diente Taxa-4(20),11(12)-dien-5 $\alpha$ ,13 $\alpha$ -diol nicht als Substrat. Ungefähr 15 % aller Metabolite induzierter *Taxus*-Zellen sind an der C14-Position oxygeniert.<sup>[106]</sup> Taxol selbst allerdings ist an dieser Stelle nicht funktionalisiert, sodass die gefundene 14 $\beta$ -Hydroxylase nicht Teil des gesuchten Taxolbiosynthesewegs ist. Allerdings ist Taxol bloß eins von vielen taxoiden Produkten, von denen es nur einen geringen Teil ausmacht. Daher ist auch die Identifizierung von Stoffwechselwegen alternativer Taxoide interessant, um ein

allgemeines Verständnis der terpenoiden Sekundärstoffbiosynthese zu gewinnen. Die Modifizierung dieser alternativen Taxoide, um den pflanzeigenen Stoffwechselfluss zu Gunsten einer verstärkten Taxolbiosynthese verändert zu können, indem die entsprechenden Enzyme manipuliert werden,<sup>[107]</sup> liegt allerdings wohl noch in ferner Zukunft.

Die von Jennewein und Mitarbeitern beobachtete Verzweigung des frühen Taxoidbiosynthesewegs in 14 $\beta$ -Hydroxytaxoide und 13 $\alpha$ -Hydroxytaxoide (inklusive Taxol) konnte von Ketchum und Mitarbeitern bestätigt werden.<sup>[108]</sup> So konnten sie zeigen, dass die Acetylierung des Taxa-4(20),11(12)-dien-5 $\alpha$ -ols zu einer verstärkten Bildung von 14 $\beta$ -Hydroxytaxoiden führt. Der genaue Punkt der biosynthetischen Verzweigung in 13 $\alpha$ - und 14 $\beta$ -Hydroxytaxoiden ist zwar nicht identifiziert, wird aber nach der 10 $\beta$ -Hydroxylierung vermutet, da sowohl Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ -diol-5-acetat, als auch Taxa-4(20),11(12)-dien-5 $\alpha$ ,10 $\beta$ -diol in beide Formen der Hydroxytaxoide eingebaut werden können.<sup>[108]</sup>

Die Suche nach den noch fehlenden Hydroxylasen gestaltete sich zunehmend schwieriger, da ihre Reihenfolge und damit die entstehenden Intermediate als weitere Substrate nicht bekannt waren. Stellvertretend wurde daher Taxusin als Substrat eingesetzt, um die fehlenden C1-, C2- und C7-Hydroxylasen zu finden. Dazu wurde die aus induzierten *Taxus*-Zellen erstellte cDNS-Bibliothek in einem geeigneten Hefestamm exprimiert, der die für eine P450-Aktivität nötige Cytochrom-P450:NADPH-Reduktase koexprimierte.<sup>[109]</sup> Durch die Zufütterung des taxoiden Substrats *in vivo* konnten sowohl die 7 $\beta$ -Hydroxylase,<sup>[110]</sup> als auch die 2 $\alpha$ -Hydroxylase<sup>[111]</sup> identifiziert werden.

Neben der von der 5-*O*-Acetyltransferase katalysierten Acetylierung an der C5-Position<sup>[100]</sup> wird Taxol an vier weiteren Positionen durch Acyltransferasen funktionalisiert. Um die entsprechenden Enzyme des Taxolbiosynthesewegs zu finden, wurde eine cDNS-Bibliothek einer induzierten *Taxus*-Zellkultur als Matrize für die Amplifikationsreaktion mit pflanzlichen transacylasespezifischen, degenerierten Primern verwendet. Im Zuge dessen wurde ein PCR-Fragment gefunden, welches auf Nukleotidebene eine 80 %ige Homologie zu der bereits identifizierten Taxadien-5 $\alpha$ -ol-*O*-acetyltransferase aufwies. Durch Hybridisierung mit der cDNS-Bibliothek wurde der Vollängklon erhalten und heterolog exprimiert. Der Zellaufschluss wurde mit 10-Deacetylbaccatin III und radioaktiv markiertem Acetyl-CoA inkubiert. Anhand des entstandenden Produkts Baccatin III konnte das Enzym als 10-Deacetylbaccatin III-10-*O*-acetyltransferase identifiziert werden.<sup>[112]</sup> Auch die 2 $\alpha$ -*O*-Benzoyltransferase wurde durch diese Methodik entdeckt. Als Substrat diente hier das semisynthetische 2-Debenzoyl-7,13-diacetylbaccatin III.<sup>[113]</sup>

Die Wirkung des Cytostatikums Taxol basiert bekanntermaßen auf der C13-Seitenkette des Moleküls,<sup>[114]</sup> sodass die in die entsprechende Transferasereaktion der Seitenkette involvierten Enzyme von besonderer Bedeutung sind. Durch die funktionale Durchmusterung einer cDNS-Bibliothek<sup>[100]</sup> konnte ein Klon identifiziert werden, der für die Taxoid C13-*O*-phenylpropanoyltransferase kodiert. Dieses Enzym katalysiert mit  $\beta$ -Phenylalanonyl-CoA als Acyldonor die 13-*O*-Acylierung von Baccatin III zu *N*-Debenzoyl-2'-deoxytaxol.<sup>[115]</sup> Die fehlende 3'-*N*-Debenzoyl-2'-deoxytaxol-*N*-benzoyltransferase wurde praktisch gleichzeitig aufgrund der von ihr katalysierten stereoselektiven Umwandlung von *N*-Debenzoyl-(3'*RS*)-2'-deoxytaxol mit Benzoyl-CoA zu 2'-Deoxytaxol gefunden.<sup>[116]</sup> Das für die 2'-Hydroxylierung der Seitenkette kodierende Gen wurde bisher nicht identifiziert. Allerdings konnte mit *Taxus*-Mikrosomen gezeigt werden, dass die Reaktion durch ein Cytochrom-P450-Enzym katalysiert wird und erst nach der Acylierung von Baccatin III stattfindet.<sup>[117]</sup> Die genaue Reihenfolge der Reaktionen von  $\beta$ -Phenylalanoylbaccatin III bis hin zum Taxol wurde von Croteau und Mitarbeitern aufgeklärt. Diese zeigten, dass die bereits identifizierte *N*-Benzoyltransferase bevorzugt *N*-Debenzoyltaxol und nicht *N*-Debenzoyl-2'-deoxytaxol umsetzt, die 2'-Hydroxylierung also vor der *N*-Benzoylierung stattfindet.<sup>[118]</sup> Weiterhin wurde die *N*-Benzoyltransferase auf ihre Selektivität für das Cosubstrat Benzoyl-CoA untersucht. Da weder Hexanoyl-CoA noch Butanoyl-CoA transferiert wurden, ist davon auszugehen, dass weitere *N*-Benzoyltransferasen für die Synthese verwandter *N*-acylierter Taxoide existieren.<sup>[118]</sup> Der Einbau der Seitenkette des Taxolmoleküls ist in Abbildung I.16 zusammengefasst.

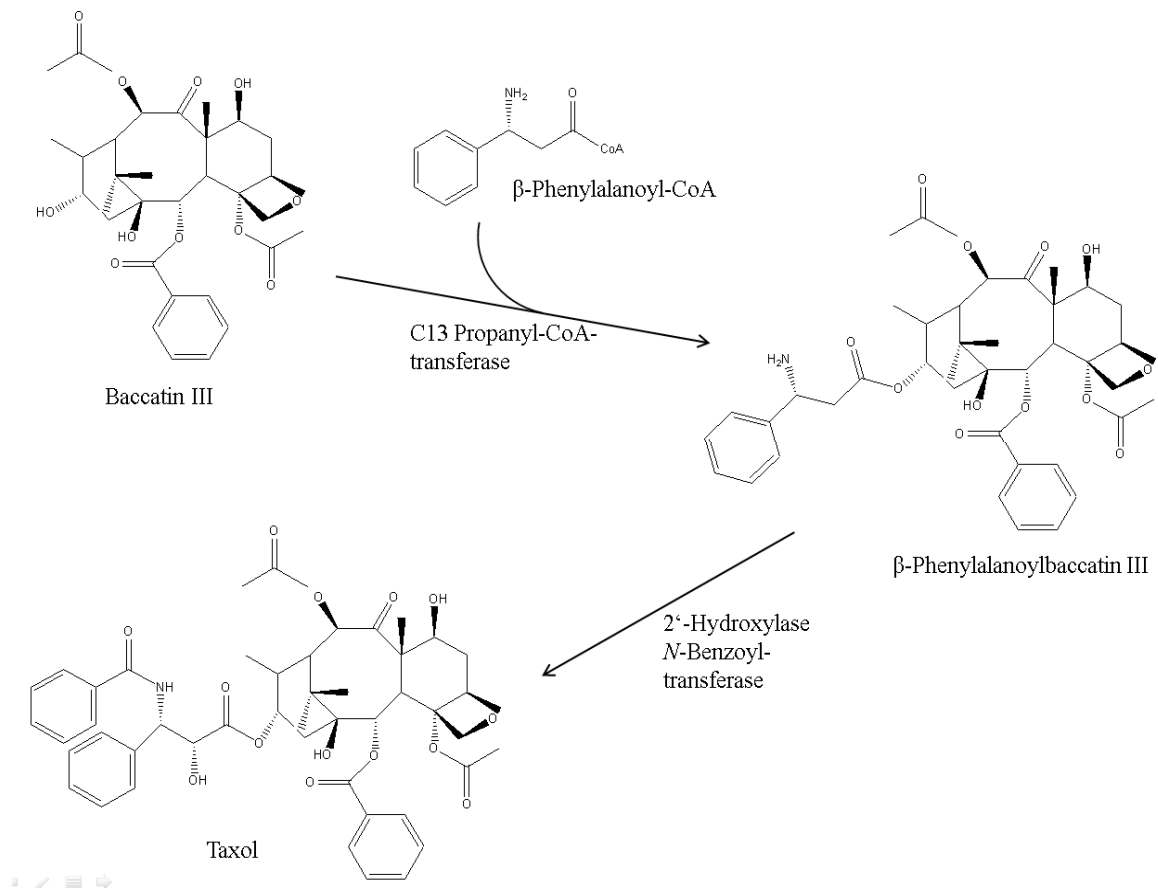
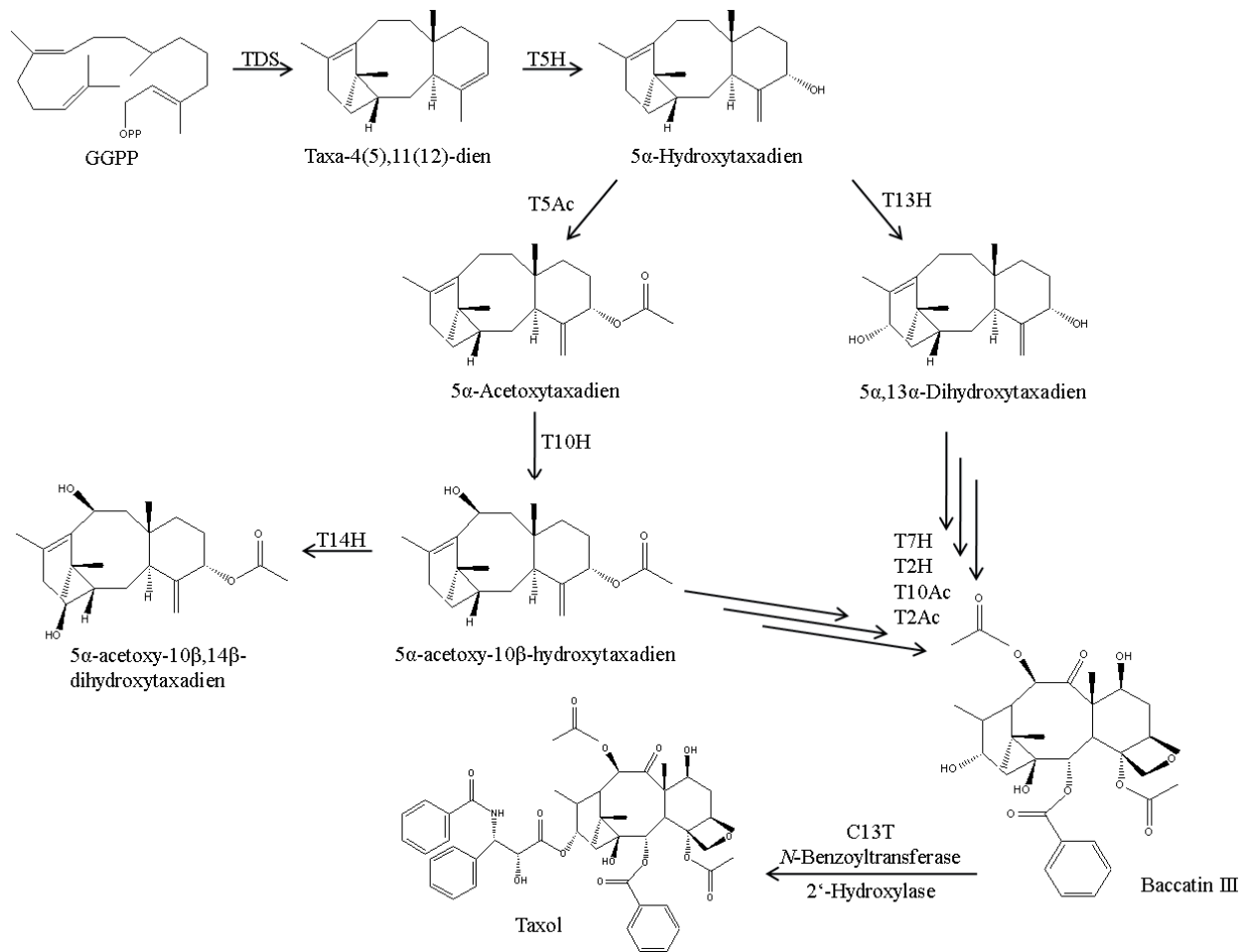


Abbildung I.16: Einbau der Seitenkette zur Biosynthese von Taxol <sup>[115]</sup> <sup>[116]</sup>

Wie auch bei den in die Taxoidbiosynthese involvierten Cytochrom-P450-Monooxygenasen, die mit 70 % eine hohe Ähnlichkeit zueinander und weit weniger zu anderen pflanzlichen P450-Monooxygenasen aufweisen,<sup>[119]</sup> zeigen auch die taxoiden Acetyltransferasen eine Ähnlichkeit von über 65 % zueinander.<sup>[93]</sup> Eine Zusammenfassung aller bekannten Enzyme aus dem Taxolbiosyntheseweg ist in Abbildung I.17 dargestellt.



**Abbildung I.17: Schematisch dargestellter Taxolbiosyntheseweg.** Aufgezeigt sind alle identifizierten Enzyme. **GGPP:** Geranylgeranyldiphosphat; **TDS:** Taxadiensynthase; **T5H:** Taxadien-5α-hydroxylase; **T5Ac:** Taxadien-5α-O-acetyltransferase; **T13H:** Taxan-13α-hydroxylase; **T10H:** Taxan-10β-hydroxylase; **T14H:** Taxan-14β-hydroxylase; **T7H:** Taxoid-7β-hydroxylase; **T2H:** Taxoid-2α-hydroxylase; **T10Ac:** 10-Deacetylbaaccatin III-10-O-acetyltransferase; **T2Ac:** 2α-O-Benzoyltransferase; **C13T:** C13-O-Phenylpropanoyltransferase

### I.3 Metabolic engineering

*Metabolic engineering* wird definiert als „die Verbesserung zellulärer Aktivität durch die Manipulation von enzymatischen, regulatorischen und Transportfunktionen der Zelle durch die Verwendung rekombinanter DNS-Techniken“.<sup>[120]</sup> Durch den Einsatz rekombinanter DNS-Technologie wird die Integration neuer Enzyme oder ganzer Stoffwechselwege ermöglicht und erlaubt durch die Veränderung verschiedener Produktionsorganismen die Nutzung unterschiedlicher Energiequellen und Startmaterialien. Allerdings verkörpert der zelluläre Stoffwechsel ein komplexes Netzwerk, dessen Funktionalität auf diversen Regulationsmechanismen basiert. Die Konstruktion, Umleitung und Manipulation des Metabolismus durch Veränderung endogener

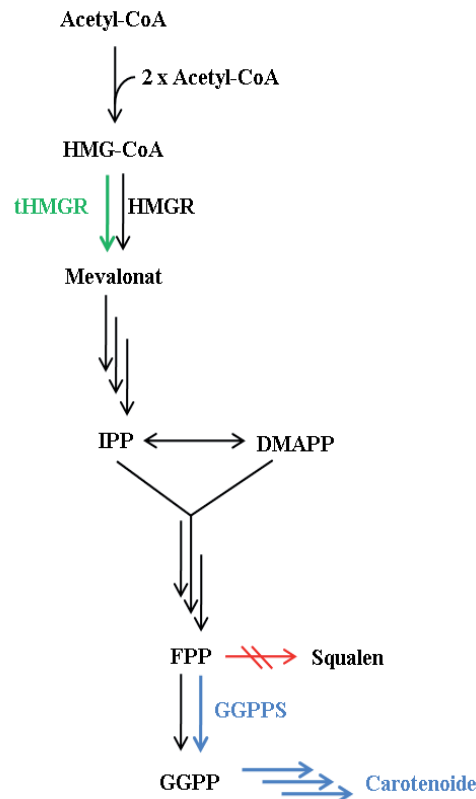


Enzymaktivitäten oder durch heterologe Enzymaktivität für eine erfolgreiche Biosynthese gewünschter Verbindungen ist daher abhängig von einem grundsätzlichen Verständnis der zellulären Vorgänge. Viele der benötigten Informationen wurden bereits anhand von Genomsequenzierungsprojekten oder *metabolic flux*-Analysen verschiedener Modellorganismen erhalten, sodass *metabolic engineering* mittlerweile sowohl für die Herstellung von Bulk- und Feinchemikalien, als auch von alternativen Energiequellen wie Biokraftstoffen Verwendung findet.<sup>[121]</sup>

### ***I.3.1 Metabolic engineering von Mikroorganismen zur Produktion von Terpenen***

Zahlreiche Strategien zur rekombinanten Produktion von Terpenen wurden bereits entwickelt. Neben der heterologen Expression entsprechender Terpenbiosynthesegene basieren auch diese *metabolic engineering*-Ansätze häufig auf einer Optimierung des Mikroorganismus-eigenen Primärmetabolismus. Dadurch soll eine ausreichende Bereitstellung der benötigten isoprenoiden Vorstufen DMAPP und IPP, bzw. ihrer effizienten Kondensation zu den entsprechenden Terpenvorstufen gewährleistet werden. Die im Folgenden beschriebenen Manipulationen sind in Abbildung I.18 zusammengefasst. Sie wurden hauptsächlich anhand der Carotinoidbiosynthese durchgeführt, da Veränderungen in Produktspektrum und -stärke aufgrund deren Farbgebung leicht zu detektieren sind.<sup>[122]</sup> So wurden zur rekombinanten Terpenproduktion in Hefen nicht nur die für die Synthese von  $\beta$ -Carotin benötigten Gene,<sup>[123]</sup> sondern auch eine GGPP-Synthase koexprimiert,<sup>[124]</sup> um den endogenen Vorrat an benötigter Diterpenvorstufe GGPP zu erhöhen. Durch Inaktivierung des Squalensynthasegens wurde dieser Vorrat zusätzlich gesteigert, da der Kohlenstofffluss zu ungewollten Sesquiterpenen verhindert wurde.<sup>[125]</sup> Die weitere Modifizierung des Mevalonatwegs (I.2.1.1), über den in Hefen die Vorstufen aller Isoprenoide IPP und DMAPP synthetisiert werden, resultierte erwartungsgemäß in Produktsteigerungen von sowohl rekombinanten Sesqui-<sup>[69][126]</sup> als auch Diterpenen.<sup>[125]</sup> Erreicht wurde dies durch die Expression eines verkürzten HMG-CoA-Reduktasegens, dessen Produkt ausschließlich aus der katalytischen Domäne des Enzyms besteht (I.2.1.1). Die fehlende regulatorische Domäne führt zu einer Deregulation des Enzyms, welches in voller Länge den limitierenden Schritt des Mevalonatwegs katalysiert.<sup>[63]</sup> Durch Umgehung dieser Limitierung liegen die Endprodukte IPP und DMAPP für die weitere Synthese von Terpenen in größeren Mengen vor.

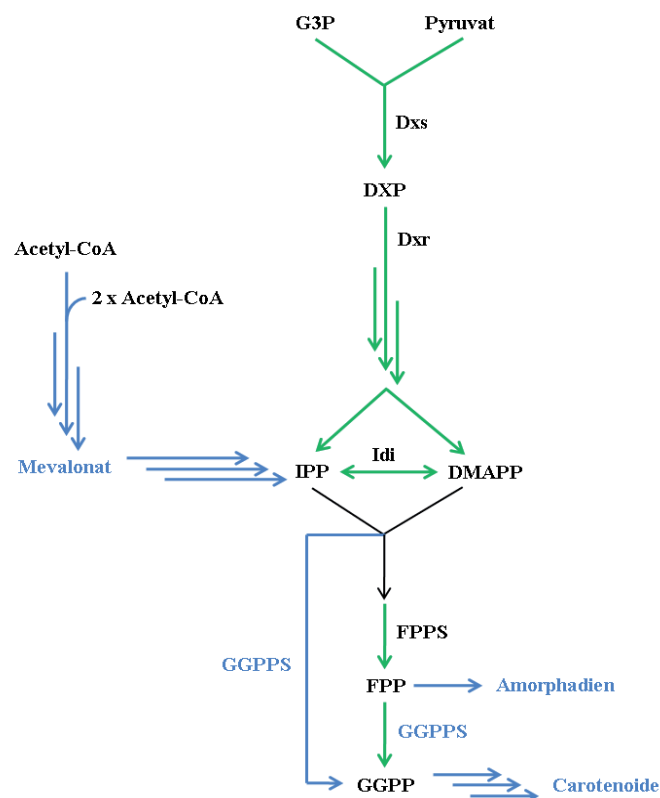




**Abbildung I.18: Schematische Darstellung der metabolischen Veränderungen von Hefe zur rekombinanten Produktion von Terpenen.** **Grün:** Homolog exprimierte Gene, **Blau:** Heterolog exprimierte Gene und die dadurch zusätzlich synthetisierten Metabolite; **Rot:** Inaktivierte Gene

Neben dem *metabolic engineering* von Hefen wurde auch *E. coli* für die rekombinante Produktion von Terpenen verwendet. Auch hier wurden für eine erfolgreiche Syntheseleistung nicht nur die entsprechenden Terpenbiosynthesegene exprimiert, sondern der bakterielle Primärstoffwechsel optimiert (Abbildung I.19). Wie schon bei *S. cerevisiae* gezeigt, führte die alleinige Expression von Carotinoidbiosynthesegenen in *E. coli* ebenfalls zu keinem detektierbaren Produkt. Bei zusätzlicher Expression der archaebakteriellen GGPP-Synthase aus *Archaeoglobus fulgidus* und Überexpression des *E. coli* eigenen IPP-Isomerasogens hingegen wurden fast 1,5 mg/g Trockengewicht des Carotinoids Astaxanthin synthetisiert.<sup>[127]</sup> Dieser Ertrag konnte durch gerichtete Evolution der GGPP-Synthase mittels *error prone* PCR verdoppelt werden.<sup>[128]</sup> Wie entscheidend die Bereitstellung von IPP und DMAPP für eine erfolgreiche Terpenproduktion ist, konnte durch verschiedene Optimierungsansätze des bakteriellen DXP-Wegs gezeigt werden. So vervierfachte sich die Carotinoidproduktion durch die Integration des Isomerasogens aus der Grünalge *Haematococcus pluvialis*.<sup>[129]</sup> Das entsprechende Gen aus anderen Organismen wie den Hefen *S. cerevisiae* oder *Phaffis rhodozyma* führte zwar ebenfalls zu einer leichten Produktionssteigerung

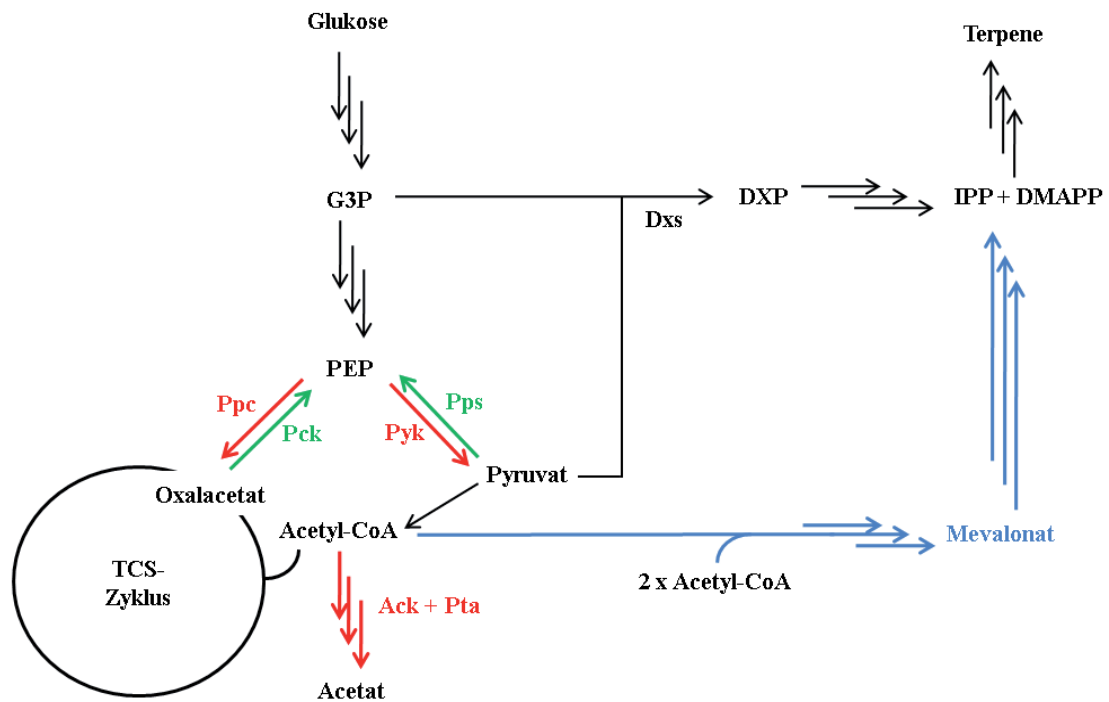
im Vergleich zum bakteriellen Gegenstück, erreichte aber nicht die hohen Konzentrationen des rekombinanten Bakterienstamms mit dem Grünalgenen.<sup>[129]</sup> Einen weiteren limitierenden Schritt des DXP-Wegs stellt die von der DXP-Synthase katalysierte Reaktion dar. Verschiedene Gruppen konnten eine bis zu 10 fache Steigerung der Diterpenproduktion allein durch die Überexpression des DXP-Synthasegens detektieren.<sup>[128][130][131]</sup> Dass die von der DXP-Synthase und IPP-Isomerase katalysierten Schritte nicht die einzig limitierenden des DXP-Wegs sind, zeigte Albrecht et al 1999. Sie integrierten sowohl die Gene für DXP-Synthase und IPP-Isomerase, als auch das Reduktoisomerasegen (I.2.1.2). Die Produktänderung durch Kombination aller drei Gene führte zu einer Diterpensteigerung, die die Summierung der jeweiligen Produktsteigerung durch Expression der einzelnen Gene ergab.<sup>[132]</sup> Die signifikanteste erzielte Produktionssteigerung erfolgte allerdings durch den Austausch aller Promotoren des DXP-Wegs mit einem starken viralen Promotor, welcher von *E. coli* RNS-Polymerasen erkannt wird. Dieser chromosomale Eingriff resultierte in einer 24 fachen Steigerung der Diterpenkonzentration.<sup>[133]</sup>



**Abbildung I.19: Schematische Darstellung metabolischer Veränderungen des *E. coli*-Stoffwechsels zur rekombinanten Produktion von Terpenen. Grün:** Homolog exprimierte Gene; **Blau:** Heterolog exprimierte Gene und dadurch entstandene zusätzliche Metaboliten

Neben der rekombinanten Produktion von Diterpenen wurde auch diejenige von Sesquiterpenen in *E. coli* untersucht. Für die Generierung des Amorpha-4,11-dien (I.1.2) produzierenden Stamms verwendeten Martin und Mitarbeiter ein kodonoptimiertes Amorphadiensynthasegen,<sup>[134]</sup> da bereits gezeigt worden war, dass pflanzliche Sesquiterpenzyklasen in *E. coli* aufgrund von Kodonpräferenzen schlecht exprimiert werden.<sup>[135]</sup> Um die intrazelluläre Bereitstellung des ADS-Substrats FPP zu erhöhen, wurden sowohl das FPP-Synthasegen aus *E. coli*, als auch die sieben Gene des Mevalonatwegs aus *S. cerevisiae* überexprimiert. Der Mevalonatweg wurde verwendet, um unbekannte Kontrollmechanismen des eigenen DXP-Wegs zu umgehen. Durch diese Modifizierungen wurde ein *E. coli*-Stamm generiert, der in der Lage war, mit 112 mg/l seine Amorpha-4,11-dien-Produktion um das 100 fache zu steigern.<sup>[134]</sup>

In *E. coli* wurde allerdings nicht nur die Optimierung des Isoprenoidbiosynthesewegs selbst analysiert, sondern auch verschiedene Modifizierungen des zentralen Stoffwechsels untersucht (Abbildung I.19). Der DXP-Weg basiert auf einem stöchiometrischen 1:1-Verhältnis von G3P und Pyruvat (Abbildung I.9). Pyruvat stellt dabei einen zentralen Metaboliten dar, der in viele Stoffwechselwege involviert ist und daher in größeren Mengen vorhanden ist. Im Gegensatz dazu wird G3P hauptsächlich durch die Glykolyse verstoffwechselt, indem es über Phosphoenylpyruvat (PEP) zu Pyruvat umgewandelt wird, und steht für die Synthese von Isoprenoiden nicht mehr zur Verfügung (Abbildung I.20). Die Umwandlung von PEP zu Pyruvat wird dabei durch die Pyruvatkinase (Pyk) katalysiert, die umgekehrte Richtung durch die PEP-Synthase (Pps). Die Überexpression des Pps-Gens und damit eine verstärkte Umwandlung von Pyruvat zu PEP in Richtung G3P resultierte daher in einer 5 fachen Steigerung der diterpenoiden Lykopenproduktion, eine Deletion von *Pyk* wiederum in einer dreifachen Lykopenakkumulation.<sup>[136]</sup>



**Abbildung I.20: Schematische Darstellung modifizierter Bereiche des Zentralstoffwechsels von *E. coli*** <sup>[136][137]</sup>  
**Rot:** Deletierte Gene; **Grün:** Überexprimierte Gene; **G3P:** Glycerinaldehyd-3-phosphat; **PEP:** Phosphoenolpyruvat; **Ppc:** PEP-Carboxylase; **Pck:** PEP-Carboxykinase; **Pps:** PEP-Synthase; **Pyk:** Pyruvatkinase; **TCS:** Tricarbonsäure; **DXP:** 1-Deoxy-D-xylulose-5-phosphat; **Dxs:** DXP-Synthase; **IPP:** Isopentenylidiphosphat; **DMAPP:** Dimethylallyldiphosphat; **Ack:** Acetatkinase; **Pta:** Phosphotransacetylase; **Blau:** Integrierter Mevalonatweg

Die limitierte Verfügbarkeit von G3P für die Terpensynthese wurde durch die Überexpression, bzw. Deletion der Gene von PEP-Carboxykinase (Pck) und PEP-Carboxylase (Ppc) bestätigt (Abbildung I.20). Die entsprechenden Enzyme Ppc katalysieren die Umwandlung von PEP zu Oxalacetat OAA (Ppc), bzw. von OAA zu PEP (Pck). So konnte bei Überexpression des Pck-Gens eine dreifache Lykopenproduktion erzielt werden, wohingegen die Überexpression des Ppc-Gens zu einer um 30 % niedrigeren Lykopenakkumulation führte.<sup>[136]</sup> Durch die Integration des Mevalonatwegs aus *Streptomyces* CL190 <sup>[138]</sup> in *E. coli* entstand eine Abhängigkeit der Biosynthese der Isoprenoidvorstufen IPP und DMAPP von der Verfügbarkeit an Acetyl-CoA. Letzteres wird durch die Enzyme Phosphotransacetylase und Acetatkinase allerdings über Acetylphosphat zu Acetat verstoffwechselt und steht für die rekombinante Biosynthese von IPP und DMAPP über den integrierten Mevalonatweg nur begrenzt zur Verfügung (Abbildung I.20). Eine Deletion der entsprechenden Gene<sup>[139]</sup> führte daher zu einer Verdoppelung der Diterpenproduktion.<sup>[137]</sup>

Um weitere Gene zu identifizieren, deren *knockout* Einfluss auf die Isoprenoidbiosynthese von *E. coli* haben könnte, wurde eine stöchiometriebasierende genomweite Fluxanalyse durchgeführt.<sup>[140]</sup> Diese sollte Aufschluss geben, welche Deletionen eine Erhöhung der Vorstufen- und Cofaktorbereitstellung ermöglichen. Im Rahmen dessen wurden die Gene für die Glutamatdehydrogenase, die Pyruvatdehydrogenase, die Phosphoglucomutase II und die Formiatdehydrogenase einzeln und in verschiedener Kombination deletiert. Nur bei Ausschalten aller Gene gleichzeitig konnte eine 40 %-Steigerung der Lykopinsynthese erzielt werden. Um weiterhin unbekannte, und daher nicht berechenbare Regulationsmechanismen zu identifizieren, wurde eine Transposon-Bibliothek generiert, die eine Deletion in folgenden Bereichen ergab: *yjfP* (nicht annotiert), *yjiD* (unbekannte Funktion) und *rssB*, dessen Genprodukt in die Proteolyse des Sigmafaktors involviert ist.<sup>[141]</sup> Die sieben Stämme mit den stöchiometrisch ermittelten Deletionen<sup>[140]</sup> wurden in den durch die Transposon-Bibliothek detektierten Genen einzeln oder gleichzeitig deletiert und die so entstandenen 64 Mutationsstämme untersucht. Während die mathematisch berechneten Deletionsziele einen mit der Deletionsanzahl linear steigende Lykopinproduktion zeigten, führten die Veränderungen der Regulationsfaktoren einzeln zu einer größeren Veränderung, in Kombination allerdings nicht immer zu einer Verbesserung.<sup>[141]</sup> Eine grundsätzliche Abhängigkeit der Terpenproduktion vom gewählten *E. coli*-Stamm, bzw. von der Promotorstärke der über- oder heterolog exprimierten Gene konnten Kim & Keasling zeigen. Eine zu starke Induktion der Gene führte aufgrund der zu hohen Syntheseleistung zu geringerem Wachstum der Zellen. Weiterhin wurden signifikant höhere Erträge in den Stämmen XL1-blue und DH5 $\alpha$  als in JM101 erzielt.<sup>[130]</sup>

### ***1.3.2 Metabolic engineering zur Produktion von Taxol***

#### ***1.3.2.1 In vivo-Produktion in Escherichia coli***

Neben der Expression einzelner Gene der Taxolbiosynthese in *E. coli* <sup>[89][110-113][115-116]</sup> zur Charakterisierung ihres Genprodukts, wurden auch Versuche zur bakteriellen *in vivo*-Produktion von Intermediaten des Taxolbiosynthesewegs unternommen.<sup>[142]</sup> Dazu wurde das Taxadiensynthasegen in *E. coli* exprimiert. Trotz nachgewiesener *in vitro*-Aktivität der gebildeten Taxdiensynthase führte dies zu keiner detektierbaren Akkumulation von Taxadien. Basierend auf ähnlichen Beobachtungen im Rahmen der rekombinanten Carotinoidproduktion in *E. coli* <sup>[127][128]</sup> wurde der Grund in einer unzureichenden endogenen Verfügbarkeit des Substrats GGPP vermutet. Um diese zu erhöhen, wurden sowohl das pflanzliche GGPP-Synthasegen, als auch das *E. coli*

eigene IPP-Isomerasegen koexprimiert. Die überexprimierte IPP-Isomerase sollte das für die Synthese von GGPP benötigte 3:1-Verhältnis der Vorstufen IPP und DMAPP gewährleisten, die über den DXP-Weg in einem Verhältnis von 1:1 gebildet werden (Abbildung I.9). Dadurch konnte 0,5 mg/l Taxadien nachgewiesen werden.<sup>[142]</sup> Die Expression von *dxs* in *E. coli* führt bekanntermaßen zu einer erhöhten Isoprenoidsynthese.<sup>[128][130-131][143]</sup> Dieses Gen wurde daher auch von Huang und Mitarbeitern verwendet, um die rekombinante Taxadienproduktion zu steigern. Die zusätzliche homologe Expression des DXP-Synthasegens führte zu einer 2,6 fachen Produktsteigerung auf 1,3 mg/l.

Der zweite Schritt der Taxolbiosynthese, eine Hydroxylierung von Taxadien an der C5-Position, wird von einem Cytochrom-P450-Enzym katalysiert.<sup>[98]</sup> Wie in 0 beschrieben, sind eukaryotische P450 Monooxygenasen durch eine N-terminale Transmembranhelix in der ER-Membran verankert<sup>[97]</sup> und funktionell abhängig von einer NADPH-Reduktase, die den benötigten Elektronentransfer von NADPH in ihr katalytisches Zentrum gewährleistet.<sup>[96]</sup> Für die Expression in *E. coli*, ein Bakterium, das selbst keine P450-Enzyme aufweist, müssen eukaryotische Cytochrom-P450 daher nicht nur N-terminal modifiziert, sondern auch mit der entsprechenden Reduktase fusioniert werden. Zwar wurden vereinzelt bereits funktionelle Fusionsproteine von Cytochrom-P450-Monooxygenasen mit der Reduktase in *E. coli* exprimiert,<sup>[144][145]</sup> aber eine erfolgreiche Hydroxylierung von Terpenen konnte nicht gezeigt werden.<sup>[146]</sup> Im Gegensatz zur semisynthetischen Herstellung des Sesquiterpens Artemisinin, im Zuge derer biosynthetisches, nicht-funktionalisiertes Amorpha-4,11-dien extrahiert und chemisch in das Endprodukt Artemisinin umgewandelt wird (I.1.2), sind die in die Semisynthese von Taxol oder seinen Derivaten eingesetzten Verbindungen hochfunktionalisiert und basieren auf Cytochrom-P450-Monooxygenase-katalysierten Reaktionen. Weder die rekombinante Produktion von Taxol selbst, noch diejenige isolierbarer Intermediate für den Einsatz in die Semisynthese, in *E. coli* stellt daher ein effizientes System dar.

#### I.3.2.2 *In vivo-Produktion von Taxadien in Planta*

Da es sich bei Taxol um eine hauptsächlich aus Pflanzen isolierte Verbindung handelt, lag die Vermutung nahe, der pflanzliche Modellorganismus *Arabidopsis thaliana* sei ein optimales Expressionssystem für Taxolbiosynthesegene. Pflanzen stellen des Weiteren die Organismen mit der größten Menge an synthetisierten Terpenoiden dar, sodass die benötigten Vorstufen aus dem Primärmetabolismus auch für die rekombinante Produktion von Taxol und seinen Intermediaten in

ausreichender Form vorliegen. Die konstitutive Expression des Taxadiensynthasegens aus *T. baccata* führte allerdings zu einer Taxadienausbeute von nur 20 ng/g Trockengewicht.<sup>[147]</sup> Transgene Pflanzen zeigten gleichzeitig eine Retardation in Wachstum und Blütenbildung, sowie einen gebleichten Phänotyp aufgrund reduzierten Chlorophyll- und Carotinoidgehalts. Um herauszufinden, ob das synthetisierte Taxadien für den Expressionsorganismus grundsätzlich giftig ist, oder einfach störend auf die Synthese endogener Plastidisoprenoide wie Gibberelline, Carotinoide oder der Chlorophyllseitenkette wirkt, wurde das Taxadiensynthasegen unter einem induzierbaren Promotor exprimiert. Die Induktion zu einem für das Pflanzenwachstum weniger kritischen Zeitpunkt führte zu einer Taxadienproduktion von 600 ng/g Trockengewicht und damit zu einer 30 fachen Steigerung. Eine giftige Wirkung des gebildeten Taxadiens auf *A. thaliana* konnte somit nicht bestätigt werden. Trotz dieser ersten Erfolge einer rekombinanten Produktion von Intermediaten des Taxolbiosynthesewegs in Pflanzen, entsprach die produzierte Menge nicht der in rekombinanten *E. coli*-Zellen synthetisierten. Bedenkt man weiterhin das langsamere Wachstum und die kompliziertere Generierung transgener Pflanzen, scheint dieses System keine Produktionsalternative zu sein, wenn bereits das erste Zwischenprodukt des komplizierten Taxolbiosynthesewegs in so geringen Mengen gebildet wird. Im Gegensatz zu *A. thaliana* zeigte *Physcomitrella patens* keine durch die rekombinante Produktion von Taxadien bedingte Wachstumsretardation. Die Expression des Taxadiensynthasegens in diesem System führte zu einer Taxadienakkumulation von 0,05 % an Frischgewicht.<sup>[148]</sup>

Die genetische Manipulation von *Taxus*-Zellen für eine erhöhte Taxolproduktion wird durch die niedrige Effizienz stabiler Transformationen erschwert. Zwar entwickelte die Gruppe um Croteau ein Protokoll, das die Effizienz von 0,002 auf 1% steigerte,<sup>[108]</sup> dies erfolgte allerdings anhand der genomischen Integration eines Reportergens. Die Transformation von *Taxus*-Zellen mit einem Gen aus der Taxolbiosynthese konnte daher bis dato nicht demonstriert werden.

#### I.3.2.3 Untersuchungen zur rekombinanten Taxolbiosynthese in Hefen

Alle bekannten Taxolbiosynthesegene wurden entweder direkt zur Charakterisierung ihres jeweiligen Genprodukts in *S. cerevisiae* exprimiert (I.2.3) oder aber im Nachhinein auf ihre funktionelle Exprimierbarkeit in Hefen getestet.<sup>[149]</sup> Für die Untersuchung der Cytochrom-P450-Monooxygenasen wurde dabei der Stamm Wat11 verwendet, der eine genomisch integrierte NADPH: Cytochrom-P450-Reduktase aus *Arabidopsis thaliana* exprimiert, die für den benötigten Elektronentransfer sorgt.<sup>[109]</sup> Beobachtungen ergaben, dass die natürlich Quelle der Reduktase die



Effizienz von Cytochrom-P450-katalysierten Reaktionen beeinflussen kann.<sup>[109]</sup> Daher testeten Jennewein *et al.* für die funktionelle Expression der Taxan-10 $\beta$ -Hydroxylase in *S. cerevisiae* die gleichzeitige Expression der NADPH: Cytochrom-P450-Reduktase aus *T. cuspidata*. Ein Unterschied zur *Arabidopsis*-Variante konnte allerdings nicht gezeigt werden.<sup>[150]</sup>

Die einzige rekombinante *de novo*-Produktion von Intermediaten der Taxolbiosynthese in *S. cerevisiae* wurde von DeJong durchgeführt. Der dafür generierte Hefestamm exprimierte die fünf pflanzlichen Gene für GGPPS,<sup>[102]</sup> die um 60 Aminosäuren verkürzte Taxadiensynthase<sup>[91]</sup> sowie Taxadien-5 $\alpha$ -Hydroxylase,<sup>[99]</sup> 5 $\alpha$ -O-Acetyltransferase<sup>[100]</sup> und die Taxan-10 $\beta$ -Hydroxylase.<sup>[101]</sup> Bei Kultivierung des rekombinanten Hefestamms in Komplexmedium wurde eine Taxadienproduktion von 0,7 mg/l nachgewiesen, die durch Inkubation in Selektionsmedium auf 1 mg/l gesteigert werden konnte. Das anschließende Intermediat der Taxolbiosynthese Taxadien-5 $\alpha$ -ol konnte ebenso detektiert werden und ließ sich auf 25  $\mu$ g/l quantifizieren. Weder für 5 $\alpha$ -Acetoxytaxadien noch für 5 $\alpha$ -Acetoxy-10 $\beta$ -hydroxytaxadien konnte ein Nachweis geliefert werden.<sup>[149]</sup> Allerdings stellt *S. cerevisiae* bisher das einzige System dar, mit dem ein funktionalisiertes Taxan rekombinant synthetisiert werden konnte. Des Weiteren wurden bereits verschiedene Methoden zur Optimierung des Primärmetabolismus für eine effiziente Bereitstellung der essenziellen Isoprenoidvorstufen IPP und DMAPP entwickelt, die im Zuge der Produktion taxoider Verbindungen noch keine Verwendung fanden. Bereits bei der Synthese anderer Diterpene in Hefen resultierten sie in einer signifikanten Ertragssteigerung.<sup>[69][125-126]</sup> Ihre Anwendung würde sicherlich auch im Rahmen einer rekombinanten Produktion taxoider Verbindungen zu einer verbesserten Synthese von Taxadien und Taxadien-5 $\alpha$ -ol führen und dadurch die weitere Umsetzung zu späteren Intermediaten des Taxolbiosynthesewegs erleichtern.

#### **I.4 Zielsetzung**

Manche Terpene werden als Bestandteil essenzieller Öle und Wachse von ihrer natürlichen Quelle in relativ hoher Quantität produziert. Dies gilt nicht für Taxol, welches aufgrund seiner antikanzerogenen Wirkung als Krebsmedikament weltweit eingesetzt wird. Es konnte in *Taxus* spp. und insbesondere in Endophyten nur in sehr geringen Mengen detektiert werden, sodass die Extraktion aus diesen Quellen den hohen Bedarf von 800 bis 1000 Jahrestonnen auf lange Sicht nicht decken kann. Im Gegensatz zum Antimalariamittel Artemisinin, dessen Gehalt in der



Pflanze durch klassische Züchtung von *Artemisia* spp. signifikant erhöht werden konnte,<sup>[151]</sup> ist diese Art der Optimierung aufgrund des langsamen Wachstums von *Taxus* nicht effizient. Die entwickelten Verfahren zur Totalsynthese von Taxol bestehen aufgrund seiner komplexen Struktur aus zu vielen Reaktionsschritten mit schlechter Ausbeute und sind daher zwar akademisch interessant, für eine kommerzielle Anwendung aber nicht einsetzbar. Verschiedene Methoden zur semisynthetischen Herstellung von Taxol wurden daher bereits patentiert. Sie alle basieren auf der Extraktion der in größeren Mengen produzierten und aus der Pflanze oder Pflanzenzellkultur isolierbaren Intermediate des Taxolbiosynthesewegs Baccatin III oder 10-Deacetylbaaccatin III. Durch Optimierung der Fermentationsbedingungen und Verbesserung der Extraktionsmethoden konnte der Ertrag dieser Verbindungen erhöht werden, sodass eine Semisynthese die momentan verwendete Methode zur Produktion der Generika Taxotere und Docetaxel darstellt. Allerdings basiert sie auf einer kosten- und zeitaufwendigen Induktion der Taxolbiosynthese durch Zugabe von Methyljasmonat, was gleichzeitig auch zu einer Wachstumseinschränkung von bis zu 25 % der genomisch nicht stabilen Pflanzenzellkultur führt.<sup>[59]</sup> Ungleich *A. annua*, die für eine erhöhte Artemisininsynthese bereits genetisch verändert werden konnte,<sup>[152]</sup> existieren für *Taxus*-Zellkulturen weiterhin keinerlei Methoden für einen erfolgreichen und stabilen Gentransfer, sodass in absehbarer Zukunft eine signifikante Steigerung der Taxolproduktion mithilfe von pflanzlichen Zellkulturen nicht zu erwarten ist. Die rekombinante Produktion von Taxol oder seinen Intermediaten durch *metabolic engineering* von schnellwachsenden Mikroorganismen verkörpert daher eine attraktive Alternative mit zahlreichen Vorzügen. Im Gegensatz zur *Taxus*-Zellkultur bieten Modellorganismen wie *E. coli* oder *S. cerevisiae* verschiedenste Möglichkeiten zur genetischen Modifikation, umfangreiche Informationen ihrer Regulationsmechanismen, sowie einen aufgeklärten Metabolismus. Eine rekombinante Produktion in diesen Organismen ist nicht wachstumsabhängig und basiert auf der Verstoffwechsellung günstiger Kohlenstoffquellen. Die Integration der Taxolbiosynthesegene in einen Organismus, der selber keine taxoiden Verbindungen und kaum Diterpene herstellt, führt weiterhin zu einem einzigen Produkt, das nicht aus einer Fraktion mit hunderten von anderen Taxanen (inklusive Herzgiften wie Taxinen) isoliert werden muss. Eine variierende Kombination der integrierten Gene ermöglicht überdies auch die Biosynthese eines veränderten Endprodukts. Die erfolgreiche Generierung eines Stamms für die rekombinante Herstellung eines späten Intermediats aus dem Taxolbiosyntheseweg könnte daher als „Plattform“ für die Produktion von Taxolderivaten mit verbesserten pharmakokinetischen Eigenschaften eingesetzt werden.

Alle identifizierten Gene aus dem Taxolbiosyntheseweg wurden bereits erfolgreich in *S. cerevisiae* exprimiert. Erste Versuche einer rekombinanten Produktion von Intermediaten der Taxolbiosynthese in *Saccharomyces cerevisiae* führten bereits zur erfolgreichen Herstellung von Taxadien und detektierbaren Mengen an Taxadien-5 $\alpha$ -ol. Ziel dieser Arbeit war daher die Generierung eines Hefestamms mit der Fähigkeit weitere Intermediate des frühen Taxolbiosynthesewegs zu produzieren. Das sollte nicht ausschließlich anhand der heterologen Expression der entsprechenden pflanzlichen Biosynthesegene realisiert werden, sondern durch die zusätzliche Modifizierung des hefeeigenen Primärstoffwechsels. Die für eine erfolgreiche Produktion terpenoider Verbindungen benötigten Isoprenoidvorstufen werden in Hefen über den Mevalonatweg synthetisiert, in *Saccharomyces cerevisiae* aber hauptsächlich für die Biosynthese von Ergosterol eingesetzt. Im Gegensatz zur rekombinanten Synthese von Sesquiterpenen wird diejenige von Diterpenen in *Saccharomyces cerevisiae* daher durch unzureichende Mengen an benötigten Vorstufen eingeschränkt. Dies wurde sowohl für Carotinoide, als auch für Substanzen aus dem Taxolbiosyntheseweg gezeigt.<sup>[124][149]</sup> Für die Generierung eines Taxadien und dessen Derivate produzierenden Hefestamms im Rahmen dieser Arbeit sollte aufgrund dessen auch die GGPP-Synthese optimiert werden. Nach Selektion entsprechender Gene zur Optimierung der GGPP-Produktion und Untersuchung ihrer Auswirkung auf die Akkumulation rekombinanten Taxadiens, sollten die zunächst einzeln exprimierten DNS-Abschnitte fusioniert werden. Die Koexpression der Gene auf einem Plasmid würde eine Limitierung an zur Verfügung stehenden Selektionsmarkern für die Expression zusätzlicher Gene aus dem Taxolbiosyntheseweg umgehen. Für die Möglichkeit einer stabilen Erweiterung sollte das generierte Cluster weiterhin in ein künstliches Hefechromosom (*Yeast artificial chromosome*, YAC) integriert werden, da der Einsatz eines YACs für die rekombinante Produktion von Sekundärmetaboliten eine innovative Alternative zu gängigen *metabolic engineering*-Methoden darstellt.

## II Material und Methoden

### II.1 Mikroorganismen

#### II.1.1 Verwendete Stämme, Vektoren und Oligos

Die Untersuchungen dieser Arbeit wurden mithilfe der in Tabelle II.1 aufgeführten Organismen durchgeführt.

**Tabelle II.1: Verwendete Stämme**

Stamm	Genotyp	Referenz
<i>Escherichia coli</i> TOP10	F <sup>-</sup> <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74</i> <i>recA1</i> <i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galU</i> <i>galK</i> <i>rpsL</i> (Str <sup>R</sup> ) <i>endA1</i> <i>mupG</i> $\lambda$ -	Invitrogen
<i>Escherichia coli</i> 10-beta	<i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>fhuA</i> <i>lacX74</i> <i>galK16</i> <i>galE15</i> <i>mcrA</i> <i>f80d</i> ( <i>lacZ</i> $\Delta$ M15) <i>recA1</i> <i>relA1</i> <i>endA1</i> <i>mupG</i> <i>rspL</i> <i>rph</i> <i>spoT1</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> )	New England BioLabs
<i>Escherichia coli</i> <i>dam</i> <sup>-</sup> / <i>dcm</i> <sup>-</sup>	<i>ara-14</i> <i>leuB6</i> <i>fhuA31</i> <i>lacY1</i> <i>tsx78</i> <i>glnV44</i> <i>galK2</i> <i>galT22</i> <i>mcrA</i> <i>dcm-6</i> <i>hisG4</i> <i>rfbD1</i> <i>R</i> ( <i>zgb210::Tn10</i> ) Tet <sup>S</sup> <i>endA1</i> <i>rspL136</i> (Str <sup>R</sup> ) <i>dam13::Tn9</i> (Cam <sup>R</sup> ) <i>xylA-5</i> <i>mtl-1</i> <i>thi-1</i> <i>mcrB1</i> <i>hsdR2</i>	New England BioLabs
<i>Saccharomyces cerevisiae</i> CEN.PK2-1C	MATa; <i>ura3-52</i> ; <i>trp1-289</i> ; <i>leu2-3</i> 112; <i>his3</i> $\Delta$ 1; MAL2-8 <sup>C</sup> ; SUC2	[153]

Tabelle II.2 gibt die für die Klonierung und Expression verwendeten Vektoren und Plasmide an.

**Tabelle II.2: Verwendete Vektoren und Plasmide**

Vektor	Größe [bp]	Relevanter Genotyp	Referenz
pDONR221	4762	Km <sup>R</sup> Cm <sup>R</sup> <i>ccdB</i> <i>attB1</i> <i>attB2</i>	Invitrogen
pVV214	7449	Amp <sup>R</sup> <i>URA3</i> <i>attL1</i> <i>attL2</i> Cm <sup>R</sup> <i>ccdB</i> 2 $\mu$ m	[154]
pVV200	7422	Amp <sup>R</sup> <i>TRP1</i> <i>attL1</i> <i>attL2</i> Cm <sup>R</sup> <i>ccdB</i> 2 $\mu$ m	[154]
pRS313	4967	Amp <sup>R</sup> <i>URA3</i> CEN6 ARSH4	[155]
pRS315	6018	Amp <sup>R</sup> <i>LEU2</i> CEN6 ARSH4	[155]
pENTRY-IBA10	1778	Km <sup>R</sup>	IBA BioTAGnology
pNFUSE-IBA SD1	1861	Amp <sup>R</sup>	IBA BioTAGnology
pCFUSE-IBA	1842	Amp <sup>R</sup>	IBA BioTAGnology
pENTRY-IBA20	1786	Km <sup>R</sup>	IBA BioTAGnology
pAG423	8396	Amp <sup>R</sup> <i>HIS3</i> <i>attL1</i> <i>attL2</i> <i>ccdB</i> 2 $\mu$ m	[156]
pCM183	7512	Amp <sup>R</sup> <i>TRP1</i> CEN4 ARS1	[157]
pYAC4	11454	Amp <sup>R</sup> <i>TRP1</i> <i>URA3</i> <i>HIS3</i> <i>tel</i> ARS1 CEN4	[158]

Zur Amplifikation im Rahmen einer PCR (II.4.2.1) kamen die in Tabelle II.3 aufgeführten Oligonukleotide zum Einsatz.

**Tabelle II.3: Für PCR eingesetzte Oligonukleotidprimer**

Nr	Bezeichnung	Verwendungszweck (Sequenz)
1	GW TDS-Tc for	5'-Primer zur Amplifikation von <i>TDS</i> aus <i>Taxus chinensis</i> für Gateway-Klonierung (5'- <b>ggggacaagtttgtacaaaaaagcaggctaaaaat</b> gagcagtagcactggcactagcaa ')
2	GW TDS-Tc rev	3'-Primer zur Amplifikation von <i>TDS</i> aus <i>Taxus chinensis</i> für Gateway-Klonierung (5'- <b>ggggaccactttgtacaagaaagctgggtct</b> cagtgatggtgatgtacttgattggatcaatataaac-3')
3	GW GGPPS-Tc for	5'-Primer zur Amplifikation von <i>GGPPS</i> aus <i>T. chinensis</i> für Gateway-Klonierung (5'- <b>ggggacaagtttgtacaaaaaagcaggctaaaaatg</b> agttactttgacaactattttaatg-3')
4	GW GGPPS-Tc rev	3'-Primer zur Amplifikation von <i>GGPPS</i> aus <i>T. chinensis</i> für Gateway-Klonierung (5'- <b>gggaccactttgtacaagaaagctgggtct</b> caatgatgatgatggtttgcctgaatgcaatg-3')
5	GW tHMGR1-Sc for	5'-Primer zur Amplifikation von <i>tHMGR1</i> aus <i>S. cerevisiae</i> für Gateway-Klonierung (5'- <b>ggggacaagtttgtacaaaaaagcaggctaaaaatg</b> cgccgacgaatggtgaaactgaagtcacc-3')
6	GW tHMGR1-Sc rev	3'-Primer zur Amplifikation von <i>tHMGR1</i> aus <i>S. cerevisiae</i> für Gateway-Klonierung (5'- <b>gggaccactttgtacaagaaagctgggtct</b> ta atggtgatggtgggttaaatgcaggtgacggacc-3')
7	pVV NotI for	5'-Primer zur Amplifikation von <i>tHMGR1-Sc</i> aus pVV200 für die Ligation in pRS (5'- ctagag <u>cgccgct</u> ctagaactagtggatc-3')
8	pVV SacII rev	3'-Primer zur Amplifikation von <i>tHMGR1-Sc</i> aus pVV200 für die Ligation in pRS (5'-cgtg <u>cccgccg</u> caaaagcttgcataaaagccttcg-3')
9	GW UPC2.1-Sc for	5'-Primer zur Amplifikation von <i>UPC2</i> aus <i>S. cerevisiae</i> für Gateway-Klonierung (5'- <b>ggggacaagtttgtacaaaaaagcaggctaaaaatg</b> agcgaagtcggtatcacagaatc-3')
10	GW UPC2.1-Sc rev	3'-Primer zur Amplifikation von <i>UPC2</i> aus <i>S. cerevisiae</i> für Gateway-Klonierung (5'- <b>gggaccactttgtacaagaaagctgggtc</b> gtcataacgaaaaatcagagaaattgtgtgtcatcg -3')
11	UPC2.1-Sc SOE for	5'-Primer zur Amplifikation von 3'-Ende <i>UPC2.1</i> aus <i>S. cerevisiae</i> für SOE-PCR (5'-cgaatacagtggagtggtgatgatcatatgatgc-3')
12	UPC2.1-Sc SOE rev	3'-Primer zur Amplifikation von 5'-Ende <i>UPC2.1</i> aus <i>S. cerevisiae</i> für SOE-PCR (5'-gcatcatatgcatatcaccacctccactgtattcg-3')
13	pVV NotI for	5'-Primer zur Amplifikation von <i>UPC2.1-Sc</i> aus pVV200 für die Ligation in pRS (5'- ctagag <u>cgccgct</u> ctagaactagtggatc-3')
14	pVV SacI rev	3'-Primer zur Amplifikation von <i>UPC2.1-Sc</i> aus pVV200 für die Ligation in pRS (5'-ctac <u>gagctc</u> caaagcttgcataaaagccttcg-3')
15	TDSco-80 for	5'-Primer zur Amplifikation von <i>TDSco-80</i> für SOE-PCR mit pho5tt (5'-atggatgatattccaagattgtctgctaattatca-3')
16	TDSco-pho5tt rev	3'-Primer zur Amplifikation von <i>TDSco-80</i> für SOE-PCR mit pho5tt (5'-GCGTTGTAATGAGTtcattatgatgatggtgatgaacttg-3')
15	TDSco-pho5tt for	5'- Primer zur Amplifikation von pho5tt für SOE-PCR mit <i>TDSco-80</i> (5'-CCATCATCATTAATGAactcattacaacgccagtctattg-3')

**Fett:** Gateway-Sequenzen; unterstrichen: Restriktionsschnittstellen; *kursiv:* Stargate-Sequenzen; GROSSBUCHSTABEN: Überhang für SOE-PCR; **TDS:** Taxadiensynthasegen; **GGPPS:** Geranylgeranyldiphosphatsynthasegen; **tHMGR:** verkürztes HMGcoA-Reduktasegen 1 aus *S. cerevisiae*; **UPC2:** Gen für Transkriptionsfaktor; **UPC2.1:** Gen für punktmutierte Version von UPC2; **pho5tt:** Terminator des Phosphatasegens

**Fortsetzung Tabelle II.3: Für PCR eingesetzte Oligonukleotidprimer**

Nr	Bezeichnung	Verwendungszweck (Sequenz)
16	pho5tt rev	3'-Primer zur Amplifikation von pho5tt für SOE-PCR mit <i>TDSco-80</i> (5'-gcagattttaatctttcggcaaaatttag-3')
17	adh1p for	5'-Primer zur Amplifikation von adh1p für SOE-PCR mit <i>tHMGR-Sc</i> (5'-gggtgtacaatarggacttctc-3')
18	adh1p tHMGR-Sc rev	3'-Primer zur Amplifikation von adh1p für SOE-PCR mit <i>tHMGR-Sc</i> (5'-GGCCGCCATttttgtatatgagatgttattgattg-3')
19	adh1p tHMGR-Sc for	5'-Primer zur Amplifikation von <i>tHMGR-Sc</i> für SOE-PCR mit adh1p (5'-CTATCTCATATACAAAAatggcgccgaccaattggtg-3')
20	tHMGR-Sc rev	3'-Primer zur Amplifikation von <i>tHMGR-Sc</i> für SOE-PCR mit adh1p (5'-ttaatggtgatggtgggatttaatgc-3')
21	gap1p for	5'-Primer zur Amplifikation von gap1p für SOE-PCR mit <i>GGPPS-Sa</i> (5'-aaagtcacatcgcccatctctcg-3')
22	gap1p GGPPS-Sa rev	3'-Primer zur Amplifikation von gap1p für SOE-PCR mit <i>GGPPS-Sa</i> (5'-CGAAGTAAGACATttttattctttttttgttctataaatg-3')
23	gap1p GGPPS-Sa for	5'-Primer zur Amplifikation von <i>GGPPS-Sa</i> für SOE-PCR mit gap1p (5'-CAAAAAAAGAAATAAAAAatgtcttacttcgacaactacttcaac-3')
24	GGPPS-Sa gcn4tt rev	3'-Primer zur Amplifikation von <i>GGPPS-Sa</i> für SOE-PCR mit gcn4tt (5'-CAACTAATTTCTTTAAatcattcgtggtggtggtgctttctc-3')
25	GGPPS-Sa gcn4tt for	5'-Primer zur Amplifikation von gcn4tt für SOE-PCR mit <i>GGPPS-Sa</i> (5'-CACCACTAATGAttaaagaattagttggcgacgctg-3')
26	gcn4tt rev	3'-Primer zur Amplifikation von gcn4tt für SOE-PCR mit <i>GGPPS-Sa</i> (5'-gtaaaattctacttaagaaaattggcataa-3')
27	Stargate GW for	5'-Primer zur Amplifikation von <i>TDSco-pho5tt</i> inklusive 5'-Gateway-Seiten für Stargate-Klonierung (5'-aatggggtaccacaagttgtacaaaaaagc-3')
28	Stargate pho5tt rev	3'-Primer zur Amplifikation von <i>TDSco-pho5tt</i> für Stargate-Klonierung (5'-tcccgcagattttaatctttcgcaaaatttag-3')
29	Stargate gap1p fpr	5'-Primer zur Amplifikation von gap1p- <i>GGPPS-Sa</i> -gcn4tt für Stargate-Klonierung (5'-aatgaaagtcacatcgcccatctctcg-3')
30	Stargate gcn4tt rev	3'-Primer zur Amplifikation von gap1p- <i>GGPPS-Sa</i> -gcn4tt für Stargate-Klonierung (5'-tcccgtaaaattctacttaagaaaattggcataaaaaag-3')
31	Stargate adh1p for	5'-Primer zur Amplifikation von adh1p- <i>tHMGR-Sc</i> für Stargate-Klonierung (5'-aatggggtgtacaatatggacttctc-3')
32	Stargate GW rev	3'-Primer zur Amplifikation von adh1p- <i>tHMGR-Sc</i> inklusive Gateway-Seiten für Stargate-Klonierung (5'-tcccgtaaaattctacttaagaaaattggcataaaaaag-3')
33	P-gap1p for	5'-Primer zur Amplifikation von gap1p- <i>GGPPS-Sa</i> -gcn4tt für die Ligation mit pho5tt- <i>TDSco-80</i> -pAG423- <i>tHMGR-Sc</i> -adh1p (5'-P-aaagtcacatcgcccatctctcg-3')

**Fett:** Gateway-Sequenzen; unterstrichen: Restriktionsschnittstellen; *kursiv:* Stargate-Sequenzen; GROSSBUCHSTABEN: Überhang für SOE-PCR; *KURSIVE GROSSBUCHSTABEN:* Peptidlinker-Sequenz; **TDS:** Taxadiensynthasegen; **GGPPS-Sa:** Gen für GGPPS aus *Sulfolobus acidocaldarius*; **tHMGR:** verkürztes HMGCoA-Reduktasegen 1 aus *Saccharomyces cerevisiae*; **pho5tt:** Terminatorregion des Phosphatasegens aus *S. cerevisiae*; **adh1p:** Promotorregion des Alkoholdehydrogenasegens aus *S. cerevisiae*; **gap1p:** Promotorregion des Gens der generellen Aminosäurepermease aus *S. cerevisiae*; **gcn4tt:** Terminatorregion des Transkriptionsaktivators von Aminosäure-Biosynthesegenen aus *S. cerevisiae*; **P:** phosphoryliert

Fortsetzung Tabelle II.3: Für PCR eingesetzte Oligonukleotidprimer

Nr	Bezeichnung	Verwendungszweck (Sequenz)
34	P-gcn4tt rev	5'-Primer zur Amplifikation von gap1p- <i>GGPPS-Sa</i> -gcn4tt für die Ligation mit pho5tt- <i>TDSco-80</i> -pAG423- <i>tHMGR-Sc</i> -adh1p (5'-P-gtaaaattctacttaagaaaattggcataa-3')
35	Stargate <i>TDSco-80</i> for	5'-Primer zu Amplifikation von <i>TDSco-80</i> für Stargate-Klonierung (5'- <i>aatggatgatattccaagattgtctgctaattatca</i> -3')
36	Stargate <i>TDSco-80</i> rev	3'-Primer zu Amplifikation von <i>TDSco-80</i> für Stargate-Klonierung (5'- <i>tcctcattatgatgatggatgaacttg</i> -3')
35	Stargate <i>TDSco-80</i> for	5'-Primer zu Amplifikation von <i>TDSco-80</i> für Stargate-Klonierung (5'- <i>aatggatgatattccaagattgtctgctaattatca</i> -3')
36	Stargate <i>TDSco-80</i> rev	3'-Primer zu Amplifikation von <i>TDSco-80</i> für Stargate-Klonierung (5'- <i>tcctcattatgatgatggatgaacttg</i> -3')
37	Stargate GGPPS-Sa for	5'-Primer zu Amplifikation von <i>GGPPS-Sa</i> -gcn4tt-adh1p- <i>tHMGR-Sc</i> für Stargate-Klonierung (5'- <i>aatgatgtcttactcgacaactactcaac</i> -3')
38	<i>GGPPS-Sa</i> rev	3'-Primer zur Amplifikation von <i>GGPPS-Sa</i> -gcn4tt-pho5tt- <i>TDSco-80</i> -pAG423-cyc1tt- <i>tHMGRE-Sc</i> (5'- <i>tcattcggtggtggtgctttcttc</i> -3')
39	<i>tHMGR-Sc</i> for	5'-Primer zur Amplifikation von <i>GGPPS-Sa</i> -gcn4tt-pho5tt- <i>TDSco-80</i> -pAG423-cyc1tt- <i>tHMGRE-Sc</i> (5'- <i>atggcgccgaccaattggtg</i> -3')
40	P-pgkp for	5'-Primer zur Amplifikation von pgkp für SOE-PCR mit tefp (5'-P-gttttatatttgttgaataaagtagataattac-3')
41	pgkp-tefp rev	3'-Primer zur Amplifikation von pgkp für SOE-PCR mit tefp (5'-CATTGTTTATACGCTATTATCAGCCAgtaattgcttttagttgtatttttagtg-3')
42	pgkp-tefp for	5'-Primer zur Amplifikation von tefp für SOE-PCR mit pgkp (5'-CACTAAAAATACACAATAAAGCAATTACtggtgataatagcgtataacaatg-3')
43	P-tefp rev	3'-Primer zur Amplifikation von tefp für SOE-PCR mit pgkp (5'-P-actggtgtagaagcctttgtacagg-3')
44	GW <i>TDSco-80</i> for	5'-Primer zur Amplifikation von <i>TDSco-80</i> für SOE-PCR mit <i>GGPPS-Sa</i> über Linker (5'- <b>ggtaccacaagttgtacaaaaaagcaggctaaaaa</b> atggatgatattccaagattgtctgctaattatca-3')
45	<i>TDSco-L-GGPPS</i> rev	3'-Primer zur Amplifikation von <i>TDSco-80</i> für SOE-PCR mit <i>GGPPS-Sa</i> -gcn4tt-adh1p- <i>tHMGR</i> (5'-GTAACATCATAGAACCACCAACCatgatgatggatgaactgaattg-3')
46	<i>TDSco-L-GGPPS</i> -for	5'-Primer zur Amplifikation von <i>GGPPS-Sa</i> -gcn4tt-adh1p- <i>tHMGR</i> für SOE-PCR mit <i>TDSco-80</i> (5'-CCATCATCATGGTGGTGGTCTatgagtactttgacaactattttaag-3')
47	pYAC TEL for	5'-Primer zur Amplifikation des pYAC4-Grundgerüsts (5'-atctgcacgcaggatgctgctgg-3')
48	pYAC TEL rev	3'-Primer zur Amplifikation des pYAC4-Grundgerüsts (5'-ccgtagtgataaataaagtcttgccg-3')
49	Linker MCS for	5'- Primer für SOE-PCR zum Zusammenbau der <i>multiple cloning site</i> (5'-P-aattCTCGGTACCGGCGCGCCAATATAGAGCTCCTCAGCTTAATGCGGC CGCTATTAATTTAAATCTGGCCGCGCCT-3')

**Fett:** Gateway-Sequenzen; unterstrichen: Restriktionsschnittstellen; *kursiv:* Stargate-Sequenzen; GROSSBUCHSTABEN: Überhang für SOE-PCR; *KURSIVE GROSSBUCHSTABEN:* Peptidlinker-Sequenz; **TDS:** Taxadiensynthasegen; **GGPPS-Sa:** Gen für GGPPS aus *Sulfolobus acidocaldarius*; **gcn4tt:** Terminatorregion des Transkriptionsaktivators von Aminosäure-Biosynthesegenen aus *S. cerevisiae*; **pgkp:** Promotorregion des Phosphoglyceratkinasogens aus *S. cerevisiae*; **tefp:** Promotorregion des Gens für Transkriptionselongationsfaktor aus *S. cerevisiae*



**Fortsetzung Tabelle II.3: Für PCR eingesetzte Oligonukleotidprimer**

Nr	Bezeichnung	Verwendungszweck (Sequenz)
50	Linker MCS rev	3'-Primer für SOE-PCR zum Zusammenbau der <i>multiple cloning site</i> (5'-P-ccggAGGCCGGCCAGATTAAATTAATAGCGGCCGCATTAAGCTGAGG AGCTCTATATTGGCGCGCCGGTACCGAG-3')
51	gdp NotI for	5'-Primer zur Amplifikation des Genclusters für die Ligation in pYAC4 (5'-ggacg <u>cgccgcg</u> ctttattcttgcatccactaaatataatg-3')
52	cyc1tt NotI rev	3'-Primer zur Amplifikation des Genclusters für die Ligation in pYAC4 (5'-ggacg <u>cgccgcg</u> cgcaattaaagccttcgagcgtccc-3')

unterstrichen: Restriktionsschnittstellen; GROSSBUCHSTABEN: Überhang für SOE-PCR; **gdp**: Promotorregion des Glycerinaldehyd-3-Phosphat Dehydrogenasgens aus *S. cerevisiae*; **cyc1tt**: Terminatorregion des Cytochrom C1-Gens

Die in Sequenzierungsreaktionen verwendeten Oligonukleotidprimer sind in Tabelle II.4 aufgelistet.

**Tabelle II.4: In die Sequenzierungen eingesetzte Oligonukleotidprimer**

Bezeichnung	Verwendungszweck (Sequenz)
GW for	Primer zur Sequenzierung aus der Gateway-Seite heraus ins 5'-Ende des jeweiligen Inserts (5'-ggtaccacaagttgtacaaaaagcagc-3')
GW rev	Primer zur Sequenzierung aus der Gateway-Seite heraus ins 3'-Ende des jeweiligen Inserts (5'-gggaccactttgtacaagaaagctgggtc-3')
TDS seq int I	Primer zur Sequenzierung von <i>TDS-Tc</i> (5'-gccggattcctccaccatctt-3')
TDS seq int II	Primer zur Sequenzierung von <i>TDS-Tc</i> (5'-cgcttctcgtgttgataacattg-3')
TDS seq int III	Primer zur Sequenzierung von <i>TDS-Tc</i> (5'-gccccctttcactccaatgtctgtag-3')
TDS seq int IV	Primer zur Sequenzierung von <i>TDS-Tc</i> (5'-ggtggaaggaatccggcatggcag-3')
GGPPS seq for	Primer zur Sequenzierung von <i>GGPPS-Tc</i> (5'-gctgttcttcaggggacgc-3')
GGPPS seq rev	Primer zur Sequenzierung von <i>GGPPS-Tc</i> (5'-gctgttagccacagcaatatgc-3')
tHMGR seq for	Primer zur Sequenzierung von <i>tHMGR-Sc</i> (5'-cttgtctagcaggagatttac-3')
tHMGR seq rev	Primer zur Sequenzierung von <i>tHMGR-Sc</i> (5'-caacctccatatcttcccagcc-3')
UPC2.1 seq for I	Primer zur Sequenzierung von <i>UPC2.1-Sc</i> (5'-cgggatgttactatgtaacagg-3')
UPC2.1 seq for II	Primer zur Sequenzierung von <i>UPC2.1-Sc</i> (5'-cgccctctgcagtgttgaaatg-3')
UPC2.1 seq rev	Primer zur Sequenzierung von <i>UPC2.1-Sc</i> (5'-ctgttgcaattgcagctgctgg-3')

**TDS-Tc**: Taxadiensynthasegen aus *Taxus chinensis*; **GGPPS-Tc**: Gen für GGPPS aus *Taxus chinensis*; **tHMGR**: verkürztes HMGoA-Reduktasegen 1 aus *Saccharomyces cerevisiae*

**Fortsetzung Tabelle II.4: In die Sequenzierungen eingesetzte Oligonukleotidprimer**

<b>Bezeichnung</b>	<b>Verwendungszweck (Sequenz)</b>
GGPPS-Sa seq for	Primer zur Sequenzierung von <i>GGPPS-Sa</i> (5'-ggagcaggaataccttgacatg-3')
GGPPS-Sa seq rev	Primer zur Sequenzierung von <i>GGPPS-Sa</i> (5'-agcaccagcaataagtgcgcc-3')
pRS seq for	Primer zur Sequenzierung übers 5'-Ende von Inserts in pRS313 und pRS315 (5'-cgacggccagtgcgcgcg-3')
pRS seq rev	Primer zur Sequenzierung übers 3'-Ende von Inserts in pRS313 und pRS315 (5'-ccatgattacccaagcgcgcg-3')
seq TDSco for I	Primer zur Sequenzierung 5'-Ende von <i>TDSco</i> (5'-gccatatgattgccattcattaag-3')
seq TDSco for II	Primer zur Sequenzierung <i>TDSco</i> intern (5'-gcatggttacaatgtttcttcagatg-3')
seq TDSco for III	Primer zur Sequenzierung <i>TDSco</i> intern (5'-gctactttggatgaattgaagct-3')
seq TDSco for IV	Primer zur Sequenzierung von 3'-Ende <i>TDSco</i> über pho5tt bis gap1p (5'-gttcattgatggttatggtatgc-3')
seq gap1p for	Primer zur Sequenzierung gap1p intern über 5'-Ende von <i>GGPPS-Sa</i> (5'-gtagaaaagaacagaccaagac-3')
seq GGPPS-Sa I	Primer zur Sequenzierung über 3'-Ende von <i>GGPPS-Sa</i> bis gcn4tt (5'-ggagcaggaataccttgacatg-3')
seq gcn4tt for	Primer zur Sequenzierung von gcn4tt bis adh1p (5'-ggcgaacgctgatttcattacc-3')
seq adh1p for	Primer zur Sequenzierung von adh1p bis über 5'-Ende von <i>tHMGR-Sc</i> (5'-cttcattcacgcactactctc-3')
seq tHMGR for	Primer zur Sequenzierung von <i>tHMGR-Sc</i> intern (5'-gtgatactacgagagcggttc-3')

**UPC2.1:** Gen für punktmutierte Version des Transkriptionsfaktors UPC2 aus *S. cerevisiae*; **GGPPS:** Geranylgeranyldiphosphatsynthasegen; **TDSco:** Für die Expression in *S. cerevisiae* codonoptimiertes Taxadiensynthasegen aus *T. chinensis*; **gap1p:** Promotorregion des Gens der generellen Aminosäurepermease aus *S. cerevisiae*; **gcn4tt:** Terminatorregion des Transkriptionsaktivators von Aminosäure-Biosynthesegenen aus *S. cerevisiae*; **adh1p:** Promotorregion des Alkoholdehydrogenasegens aus *S. cerevisiae*; **tHMGR:** verkürztes HMGcoA-Reduktasegen 1 aus *Saccharomyces cerevisiae*

Für den Nachweis einer aktiven Transkription von Genen wurden die in Tabelle II.5 aufgelisteten Oligonukleotide radioaktiv markiert und als Sonde in die Northernblotanalyse eingesetzt.



**Tabelle II.5: In die Northernblotanalyse eingesetzte Oligonukleotide**

Bezeichnung	Verwendungszweck (Sequenz)
Oligo TDS-Tc	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>TDS-Tc</i> –mRNS (5'-gggtgagcttgtgaaagatgatgttgtgagaaagtgcactatccctcaaatatg-3')
Oligo GGPPS-Tc	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>GGPPS-Tc</i> –mRNS (5'-cctcttacctccagatgtaaacaatgatagagcttcataatagtttaggaacatctcc-3')
Oligo tHMGR-Sc	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>tHMGR-Sc</i> –mRNS (5'-cggagacaacctccatattctccagccatactcttctaccatttgccttaagag-3')
Oligo UPC2.1-Sc	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>UPC2.1-Sc</i> –mRNS (5'-cttttctcttcttctgggtttgtcaccgcttctgtgattctgtatataccgacttc-3')
Oligo GGPPS-sa	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>GGPPS-Sa</i> –mRNS (5'-gcagcagcttttctagagatcatgtccaagtattctgttccttgatgtcgattctgtcttcg-3')
Oligo TDSco	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>TDSco</i> –mRNS (5'-cccaaatgttcaatgttatcaaccaagacaatcttccaacaatcaatagatcatgc-3')
Oligo GGPPSco	Oligonukleotid, radioaktiv markiert eingesetzt zur Detektion von <i>GGPPSco</i> –mRNS (5'-gcagcagcttttctagagatcatgtccaagtattctgttccttgatgtcgattctgtcttcg-3')

**TDS-Tc:** Taxadiensynthasegen aus *Taxus chinensis*; **GGPPS-Tc:** Geranylgeranyldiphosphatsynthasegen aus *Taxus chinensis*; **tHMGR-Sc:** verkürztes HMG-CoA-Reduktasegen 1 aus *S. cerevisiae*; **UPC2.1:** Gen für punktmutierte Version des Transkriptionsfaktors UPC2 aus *S. cerevisiae*; **GGPPS-Sa:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **TDSco:** Für die Expression in *S. cerevisiae* codonoptimiertes Taxadiensynthasegen aus *T. chinensis*; **GGPPSco:** Für die Expression in *S. cerevisiae* codonoptimiertes GGPP-Synthasegen aus *Sulfolobus acidocaldarius*

## II.2 Nährmedien

Die verschiedenen Medien wurden mit doppelt destilliertem H<sub>2</sub>O aus der MilliQ-Anlage (dd H<sub>2</sub>O) hergestellt und vor Verwendung für 20 Minuten bei 121°C autoklaviert. Zur Herstellung von Festmedien wurden vorher 1,8% w/v Agar zugegeben.

### II.2.1 2YT-Medium

NaCl	0,5% w/v
Hefeextrakt	1,0% w/v
Trypton	1,6% w/v
pH7,5	

**II.2.2 SOC-Medium**

NaCl	0,058% w/v
Hefeextrakt	0,500% w/v
Trypton	2,000% w/v
KCl	0,019% w/v
pH7	
Nach dem Autoklavieren:	
MgCl <sub>2</sub> x 6 H <sub>2</sub> O	0,20% w/v
MgSO <sub>4</sub>	0,12% w/v
Glukose	0,36% w/v

**II.2.3 YPD-Medium**

Glucose	2% w/v
Trypton	2% w/v
Hefeextrakt	1% w/v
pH5,5	

**II.2.4 SC-Medium**

Glucose	2,00% w/v
Ammoniumsulfat	0,50% w/v
Yeast nitrogen base	0,17% w/v
pH5,5	

In Abhängigkeit der Auxotrophie eines Stammes (II.1.1) wurden vor dem Autoklavieren jeweils 40 mg/l der benötigten Aminosäuren zugegeben.

**II.2.5 Medium 88 (DSM)**

KH <sub>2</sub> PO <sub>4</sub>	0,28% w/v
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1,30% w/v
MgSO <sub>4</sub> *7H <sub>2</sub> O	0,25% w/v
CaCl <sub>2</sub> *2H <sub>2</sub> O	0,07% w/v
FeCl <sub>3</sub> *6H <sub>2</sub> O	0,02% w/v
Na <sub>2</sub> B <sub>4</sub> O <sub>7</sub> *10H <sub>2</sub> O	4,50‰ w/v
MnCl*10H <sub>2</sub> O	1,80‰ w/v
ZnSO <sub>4</sub> *7H <sub>2</sub> O	0,22‰ w/v
CuCl <sub>2</sub> *7H <sub>2</sub> O	0,05‰ w/v
Na <sub>2</sub> MoO <sub>4</sub> *2H <sub>2</sub> O	0,03‰ w/v
VO <sub>2</sub> SO <sub>4</sub> *2H <sub>2</sub> O	0,03‰ w/v
CoSO <sub>4</sub>	0,01‰ w/v
pH2 mit 4M H <sub>2</sub> SO <sub>4</sub>	

### **II.2.6     *Antibiotika***

Von den Antibiotika wurde eine 1000fach konzentrierte Stammlösung angefertigt und aliquotiert bei -20°C gelagert. Vor Zugabe zum autoklavierten Medium wurde dieses auf 50°C abgekühlt.

<b>Antibiotikum</b>	<b>Stammlösung</b>	<b>Lösungsmittel</b>
Ampicillin	100 mg/ml	Ethanol
Kanamycin	50 mg/ml	Wasser

## **II.3        Handhabung von Bakterien- und Hefestämmen**

Kultivierungen in Flüssigmedium erfolgten in Reagenzgläsern oder Erlenmeyerkolben mit Schikane. Das Volumenverhältnis von Gefäß zu Flüssigkeit betrug dabei zwischen 10:1 und 5:1. Vorkulturen wurden mittels von Platte gepickter Einzelkolonien angelegt und dienten als Inokulum der Hauptkultur.

### **II.3.1     *Zellanzucht und Zellernte von Bakterienstämmen***

*Escherichia coli* -stämmen wurden bei 37°C in gegebenenfalls Antibiotika-haltigem (0) 2YT-Medium (II.2.1) auf einem Rotationsschüttler (Innova TM 4000 Shaker, New Brunswick Scientific) mit 160 Upm inkubiert. Die Zellernte von Kulturvolumina bis zu 2 ml erfolgte für 1 min in einer Tischzentrifuge (5415D, Eppendorf) bei 16100  $\times g$ . Größere Inkubationsmengen wurden in einer Multifuge (3S-R, Heraeus) bei 4500 Upm für 5 min pelletiert.

### **II.3.2     *Zellanzucht und Zellernte von Hefestämmen***

Die Kultivierung von *Saccharomyces cerevisiae* -stämmen erfolgte bei 28°C auf einem Rotationsschüttler mit 160 Upm. Die Zellernte wurde in einer Megafuge bei 2500 Upm für 7 min durchgeführt. Für Untersuchungen von Hefestämmen auf ihre Produktionsleistung wurden diese für 24 h in 50ml des jeweils benötigten SC-Mediums (II.2.4) kultiviert. Nach einem Medienwechsel zu YPD (II.2.3) unter Zugabe von 0,5 g RP C18 Silicagel (Roth) erfolgte eine weitere Inkubationszeit von 48 h. Nach Zellernte wurde das im Überstand befindliche C18-Material durch den Einsatz eines Faltenfilters aus der Flüssigkeit gewonnen und zusammen mit dem Zellpellet gefriergetrocknet.

### ***II.3.3 Fermentation von Hefezellen im 30l-Maßstab***

Die Fermentation eines Hefestammes im 30l- Maßstab wurde einmalig durchgeführt und diente der Gewinnung von biosynthetisch hergestelltem Taxadien. Dieses fand nach Aufreinigung (II.5.5) Verwendung bei der Erstellung einer Standardgeraden zur Quantifizierung gebildeter Taxadienmengen generierter Produktionsstämme. Zu diesem Zweck wurde der im Rahmen dieser Arbeit entwickelte *Saccharomyces cerevisiae* -Stamm CEN 7 (Tabelle III.4) verwendet. Alle Fermentationsarbeiten wurden von der Fraunhofer eigenen Fermentergruppe unter der Leitung von Dr. Stefan Hellwig durchgeführt. Die Extraktion des produzierten Taxadiens aus der erhaltenen Biomasse erfolgte mittels Soxhlet-Apparatur (II.5.5).

### ***II.3.4 Stammkonservierung Reinheitskontrolle***

Zur Kurzzeitlagerung wurden die Kulturen auf Selektionsplatten ausgestrichen, bei 37°C über Nacht (Bakterienstämme) oder 28°C für 48 h (Hefestämme) inkubiert und bei 4°C gelagert. Zur Langzeitkonservierung wurden die Organismen als Glyzerinsuspension bei -80°C gelagert. Hierzu wurden die Zellen unter Selektionsbedingungen angezogen und nach sichtlicher Trübung des Mediums aliquotiert mit sterilem Glycerin (20% v/v) versetzt. Die Reaktivierung eines konservierten Stammes erfolgte durch Inkubation in Selektionsmedium. Die Reinheit der Stämme wurde mikroskopisch untersucht durch regelmäßiges Ausstreichen der Kulturen auf Agarplatten unter selektiven Wachstumsbedingungen kontrolliert.

## **II.4 Molekularbiologische Methoden**

### ***II.4.1 Isolierung von Nukleinsäuren***

#### ***II.4.1.1 Isolierung von Gesamt-DNS***

Zur Isolierung von Gesamt-DNS aus Hefezellen wurde eine 10ml Übernachtskultur geerntet (II.3.2), mit 500µl sterilem ddH<sub>2</sub>O gewaschen und die Zellen in 200µl Lysis-Puffer resuspendiert. Nach Zugabe von 200µl einer PCI (Phenol-Chloroform-Isoamylalkohol 25:24:1)-Lösung und 0,3 g Glasperlen (0,2-0,5mm Durchmesser, Roth) erfolgte der Zellaufschluss durch Schütteln (Vortexer) bei 4°C für mindestens 45 min. Dem Ansatz wurden 200µl TE-Puffer hinzugefügt und die wässrige, DNS-haltige Phase durch Zentrifugieren für 5 min bei Höchstgeschwindigkeit abgetrennt und in ein frisches 1,5 ml Eppendorfgefäß überführt. Proteinverunreinigungen wurden durch 1 ml Ethanol mit anschließender Zentrifugation für 2 min und Resuspension der Nukleinsäuren in 400µl TE-Puffer

entfernt. Es folgte eine 5 minütige RNase A-Behandlung (3µl Enzym) bei 37°C mit anschließender Ethanolfällung der DNS mittels 10µl 4M Ammoniumacetat und 1 ml Ethanol. Die nach Zentrifugieren für 2 min und 16100  $xg$  pelletierte DNS wurde bei 70°C luftgetrocknet und in 50µl TE-Puffer aufgenommen.

<b>Lysis-Puffer:</b>	Triton X-100	2% w/v
	SDS	1% w/v
	NaCl	100mM
	Tris-Cl	10mM
	EDTA	1mM
	pH8,0	
<b>TE-Puffer:</b>	Tris-Cl	10mM
	EDTA	1mM
	pH8,0	

#### II.4.1.2 Isolierung von Plasmid-DNS

Die Isolierung von Plasmid-DNA wurde mithilfe des Macherey & Nagel NucleoSpin plasmid-Kits nach Herstellerangaben durchgeführt. Die Isolierung von Hefeplasmiden erfolgte aus 15 ml einer Übernachtskultur und benötigte folgenden Zwischenschritt: Nach Zugabe des A1-Puffers wurden dem Ansatz 0,3 g Glasperlen (0,2-0,5 mm Durchmesser, Roth) zugesetzt und die Zellen durch einstündiges Schütteln (Vortexer) bei 4°C aufgeschlossen.

#### II.4.1.3 Isolierung von Gesamt-RNS <sup>[159]</sup>

Alle in die Isolierung von RNS involvierten Gefäße wurden vorher autoklaviert und mit RNase-out behandelt. Des Weiteren wurden ausschließlich gestopfte Pipettenspitzen verwendet und alle Arbeiten auf Eis durchgeführt. Für die Isolierung von Gesamt-RNS aus Hefen wurden diese über Nacht in YPD (II.2.3) kultiviert und eine Zellmenge äquivalent zu einer OD<sub>600nm</sub> von 25 geerntet (II.3.2), in 270µl Natriumacetat-Puffer aufgenommen und sowohl 0,3 g Glasperlen (0,2-0,5 mm Durchmesser, Roth) als auch 30µl einer 10% SDS-Lösung zugegeben. Der Zellaufschluss erfolgte nach Beifügen von 300µl Phenol (äquilibriert mit Wasser, vorgewärmt auf 65°C) durch sechsfache Wiederholung von 1 minütigem Schütteln bei Raumtemperatur (Vortexer) im Wechsel mit 5 min Inkubation bei 65°C (Wasserbad). Nach 2 min Inkubation auf Eis wurden 300µl Phenol-Chloroform (25:1) zugegeben und die wässrige Phase nach kurzem Vortexen und Abtrennung durch Zentrifugieren bei 4°C und 16100  $xg$  für 5 min in ein frisches Eppendorfgefäß überführt. Durch

Zugabe von 0,1 Volumen an 3M Natriumacetatlösung (pH5,2) und 2,5 Volumen an Ethanol erfolgte die RNS-Fällung bei -80°C. Nach Pelletieren bei 4°C und 16100  $\times g$  für 10min und zweimaligem Waschen mit 75% Ethanol wurde die RNS bei 45°C luftgetrocknet und in 100 $\mu$ l TE-Puffer aufgenommen. Unlösliche Verschmutzungen wurden durch einen Zentrifugationsschritt bei 4°C und 16100  $\times g$  für 1 min und Überführen der löslichen RNS in ein frisches Gefäß entfernt. Da RNS leicht degradiert, wurde sie während des Gebrauchs auf Eis, zur Langzeitaufbewahrung bei -80°C gelagert.

<b>NaAc-Puffer:</b>	Na-Acetat (pH5,2)	50mM
	EDTA (pH8,0)	10mM

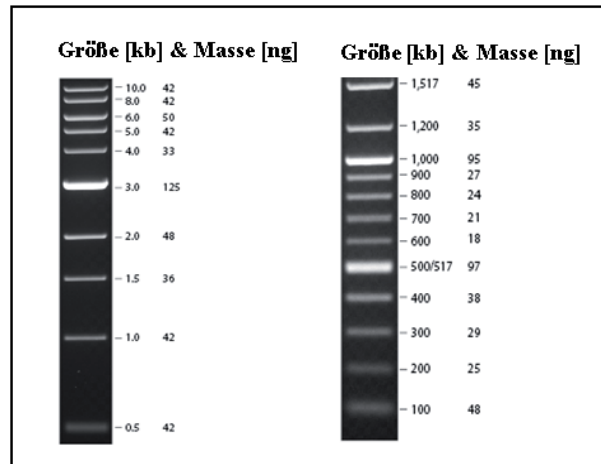
<b>TE-Puffer:</b>	Tris-Cl	5,0mM
	EDTA	0,5mM
	pH8,0	

#### *II.4.1.4 Konzentrationsbestimmung von Nukleinsäuren*

Sowohl die Konzentration als auch die Reinheit von Nukleinsäuren wurde spektrophotometrisch mithilfe des Nanodrop (ND-1000, peqlab Biotechnologie GmbH) analysiert. Das einzusetzende Volumen betrug dabei 2 $\mu$ l.

#### *II.4.1.5 Größenbestimmung von DNS-Fragmenten*

Die Größe von DNS-Fragmenten wurde routinemäßig durch Agarosegelelektrophorese bestimmt (II.4.3). Als Größenstandard dienten die 1kb-Leiter (0,5-10 kb), bzw. die 100bp-Leiter (100-1500 bp) von New England Biolabs (NEB). Abbildung II.1 zeigt die aufgetrennten Marker, sowie die jeweiligen Fragmentgrößen und -massen.



**Abbildung II.1:** Mittels Agarosegel aufgetrennte DNS-Marker, visualisiert durch Ethidiumbromid

## II.4.2 *In vitro*-Amplifikation von DNS-Fragmente

### II.4.2.1 Polymerasekettenreaktion (PCR)

Die Polymerasekettenreaktion (*Polymerase chain reaction* PCR) stellt ein Verfahren zur selektiven Vervielfältigung eines DNS-Abschnitts dar. Dabei dient die zu amplifizierende DNS als Matrize (*template*). Zwei Oligonukleotide, die komplementär (normalerweise am 5'-Ende des Matrizenstrangs) bzw. invers komplementär (am 3'-Ende des codogenen Strangs) zu dem zu vervielfältigenden DNS-Abschnitt sind, fungieren als *Primer* (Tabelle II.3). Nach Anlagerung (*annealing*) der Oligonukleotide (*Primer*) an die Matrize (*Template*) werden sie durch die Polymerase verlängert (Elongation). Die *Primer* wurden bei Invitrogen hergestellt und im Falle einer späteren Ligation (II.4.5.2) des PCR-Produkts mit entsprechenden Restriktionsschnittstellen (II.4.5.1), im Falle anschließender Gatewayklonierung (II.4.6.1) mit den benötigten Rekombinationsseiten versehen. Obwohl diverse Polymerasen mit Korrekturlesefunktion kommerziell erhältlich sind, wurden im Rahmen dieser Arbeit alle PCR mit der *Herculase II* von Stratagene durchgeführt. Diese zeichnet sich durch eine sehr hohe Amplifikationsgeschwindigkeit (1 kb / 30 sec) bei niedriger Mutationsrate aus. Die PCR-Ansätze wurden nach Herstellerangaben (Stratagene) zusammen gegeben und folgendem Amplifikationsprogramm unterzogen:

Anzahl	Funktion	Zeitintervall	Temperatur
1	Denaturierung der <i>template</i> -DNS	2 min	95°C
30	Denaturierung	30 sec	95°C
	Anlagerung der Primer	30 sec	T <sub>m</sub>
	Strangverlängerung (Elongation)	30 sec / kb	72°C
1	Finale Kettenverlängerung	5 min	72°C

**T<sub>m</sub>:** Schmelztemperatur der Primer. Diese berechnet sich aus ihrer Sequenz nach folgender Formel:  

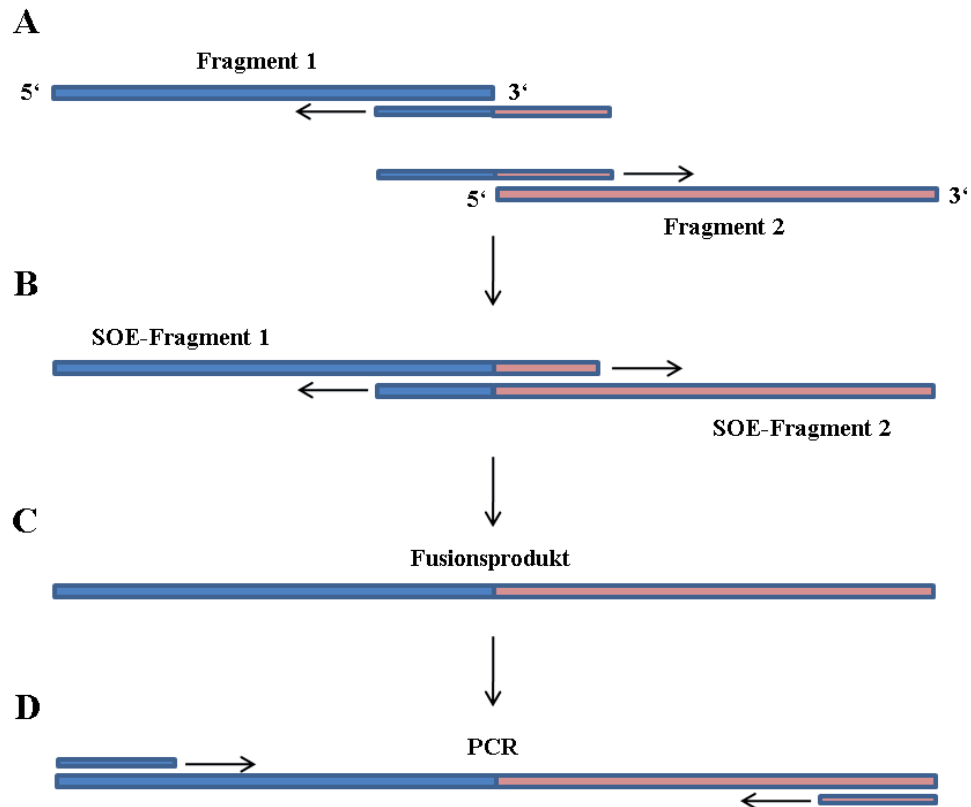
$$T_m = (A + T) * 2 + (C + G) * 4 - 6$$

Bei Produktlängen von über 6 kb, wurde die Elongationszeit auf 1min / kb verlängert. Um eine Aktivitätsdauer der Polymerase während der verdoppelten Reaktionszeit zu gewährleisten, wurden die Denaturierungsparameter auf 92 °C und 20 sec gesenkt. Die Elongationszeit wurde ebenfalls auf 20 sec erniedrigt und die –temperatur auf 68 °C gesetzt.

#### II.4.2.2 SOE-PCR

Die Methode der SOE-PCR (*Splicing by overlap extension*-PCR) wurde 1989 von HORTON *et al* entwickelt und stellt nur eine der mittlerweile zahlreichen Ansätze zur Fusionierung von DNS-Fragmenten mittels PCR dar <sup>[160]</sup>. Dabei werden die zu verbindenden DNS-Abschnitte mit Primern amplifiziert, die das später *upstream*-orientierte Fragment am 3'-Ende, das *downstream*- orientierte Fragment am 5'-Ende jeweils um eine Basenabfolge verlängern, die zueinander komplementär sind (Abbildung II.2A). Nach Denaturierung beider *templates* zu Einzelstrang-DNS, können die zu fusionierenden Fragmente während des Anlagerungsschrittes mit diesen überlappenden Enden aneinander binden (Abbildung II.2B) und anschließend durch die Polymerase in beide Richtungen zum Fusionsprodukt verlängert werden (Abbildung II.2C). Dieses Fusionsprodukt wird anschließend als Template für eine PCR eingesetzt und durch den Einsatz von Primern, die an den Enden des zuvor generierten Hybridfragments binden, vervielfältigt (Abbildung II.2D).





**Abbildung II.2:** Einzelschritte der SOE-PCR. **A:** Amplifikation der zu fusionierenden Fragmente 1 & 2 mit Primern, die je nach geplanter Position des Fragments am 5'- oder 3'- Ende einen Überhang generieren, der homolog zum Fusionspartner ist. **B:** Durch die überlappenden Überhänge können die Einzelfragmente aneinander binden und mittels PCR-Reaktion verlängert werden. **C:** Das Produkt ist ein DNS-Abschnitt, der aus Fragment 1 und Fragment 2 besteht. **D:** Dieses kann zur exponentiellen Vervielfältigung als Template in eine anschließende PCR eingesetzt werden.

Im Unterschied zu der reinen PCR (II.4.2.1), besteht die SOE-PCR also aus zwei sich abgrenzenden Reaktionen, die wie folgt durchgeführt wurden.

Anzahl	Funktion	Zeitintervall	Temperatur
1	Denaturierung der <i>template</i> -DNS	2 min	95°C
	Denaturierung	30 sec	95°C
10	Anlagerung der Primer	30 sec	50
	Strangverlängerung (Elongation)	30 sec / kb	72°C
1	Finale Kettenverlängerung	3 min	72°C
1	Denaturierung → Zugabe der Endprimer	2 min	95 °C
	Denaturierung	30 sec	95°C
25	Anlagerung der Primer	30 sec	T <sub>m</sub> Endprimer
	Strangverlängerung (Elongation)	30 sec / kb	72°C
1	Finale Kettenverlängerung	5 min	72°C

### II.4.2.3 Sequenzierung

Die Sequenzierung erfolgte nach der Kettenabbruchmethode nach Sanger.<sup>[161]</sup> Diese basiert auf dem zufälligen Einbau von ddNTPs, die keine 3'-Hydroxygruppe besitzen. Die fehlende OH-Gruppe am 3'-C-Atom verhindert eine Verknüpfung mit der Phosphatgruppe des nächsten Nukleotids, und somit auch eine weitere Verlängerung des DNS-Moleküls. Dadurch entstehen DNS-Fragmente verschiedener Länge, die statistisch bei jeder Nukleotidposition der Ursprungssequenz eines der fluoreszenzmarkierten ddNTPs eingebaut haben. Die Fragmente wurden durch Kapillarelektrophorese aufgetrennt und die per Laser detektierte Fluoreszenzemission über das "Applied Biosystems Sequencing Analysis"-Programm ausgewertet. Die Sequenzierungsreaktion wurde unter Verwendung des Thermosequenase-Kits (GE Healthcare) *in-house* von Dr. Jost Muth und Raphael Soer im Fraunhofer IME durchgeführt und die so erhaltene Sequenz anhand des DNA-Star 4.04-Programms mit den virtuell erstellten Klonierungsprodukten abgeglichen. Der dort abzugebende Sequenzierungsansatz setzte sich aus folgenden Komponenten zusammen:

Komponente	Menge
Sequenzierprimer	100pmol
DNS	800ng
ddH <sub>2</sub> O	ad 30µl

### II.4.3 Gelelektrophoretische Auftrennung von DNS

Zur gelelektrophoretischen Größenauftrennung von DNS-Fragmenten wurden die Proben 1:5 mit Probenpuffer versehen und im 0,8 % (w/v) TBE-Agarosegel bei 120V für 45 min getrennt. Als molekularer Größenstandard dienten der 1 kb-Marker, bzw. der 100 bp-Marker von NEB (II.4.1.5). Zur Visualisierung der aufgetrennten DNS wurde dem Agarosegel Ethidiumbromid (25 µg/l) beigelegt. DNS-Banden konnten bei einer Wellenlänge von 302 nm im UV Transilluminator detektiert und im Abgleich mit dem Marker ihre jeweilige Größe bestimmt werden.

<b>Probenpuffer:</b>	Bromphenolblau	0,1% (w/v)
	Xylencyanol	0,1% (w/v)
	Glyzerol	50,0% (w/v)
	TBE-Puffer 1x	ad 50ml
	pH8,3	
<b>TBE-Puffer 10x:</b>	EDTA	0,1 M
	Tris-base	2,0 M
	Essigsäure	1,0 M

## **II.4.4      *Aufreinigung von DNS***

### **II.4.4.1    *Extraktion von DNS-Fragmenten aus Agarosegelen***

Um ein DNS-Fragment aus einem Gemisch zu isolieren, wurde dieses über Agarosegelelektrophorese (II.4.3) aufgetrennt und nach Detektion der genauen Position das entsprechende Agarosestück mit einem Skalpell ausgeschnitten. Lösen der DNS aus der Agarose und weitere Aufreinigungsschritte erfolgten anhand des NucleoSpin Extract II –Kits von Macherey & Nagel (Düren) nach Herstellerangaben. Die Konzentration der isolierten DNS wurde spektrophotometrisch bestimmt.

### **II.4.4.2    *Aufreinigung von DNS über Spinsäulchen***

Um DNS nach Restriktion (II.4.5.1) oder PCR (II.2.2.1) von Verunreinigungen wie Enzymen oder dNTPS und Primern zu trennen, wurde der jeweilige Ansatz mit 250µl NT-Puffer gemischt und die DNS mit dem NucleoSpin Extract II –Kit von Macherey & Nagel (Düren) nach Herstellerangaben aufgereinigt. Auch hier erfolgte die anschließende Konzentrationsbestimmung der isolierten DNS spektrophotometrisch.

### **II.4.4.3    *Konzentrierung von DNS***

Zur Konzentrierung von DNS und Entfernung von Salzen wurde eine Polyethylenglykol- (PEG-) Fällung angewandt. Der aufzureinigende Ansatz wurde dafür auf 50 µl aufgefüllt, mit 150 µl TE-Puffer sowie 100µl PEG-Lösung versehen und für mindestens 45 min bei 16100  $\times g$  und Raumtemperatur zentrifugiert. Die pelletierte DNS wurde anschließend in der gewünschten Menge TE-Puffer gelöst und spektrophotometrisch auf ihre Konzentration untersucht.

<b>TE-Puffer:</b>	EDTA	1m M
	Tris-base	10m M
	pH8,0	
<b>PEG-Lösung:</b>	PEG 8000	30 % (w/v)
	MgCl <sub>2</sub>	30m M
	Lagerung aliquotiert bei -20°C	

## **II.4.5      *Enzymatische Modifikation von DNS***

### **II.4.5.1    *Restriktion***

Die Restriktion von DNS-Proben erfolgte durch sequenzspezifische Hydrolyse mittels Endonukleasen und diente der Modifikation von DNS-Enden zur weiteren Ligation (II.4.5.2), bzw. der Kontrolle von Ligationseignissen. Der Verdau wurde mit von New England Biolabs (NEB) bezogenen Enzymen nach Herstellerangaben durchgeführt und setzte sich wie folgt zusammen:

NEB-Puffer	3 µl
BSA (je nach Enzym)	0,3 µl
DNS	300-1000 ng
Endonuklease	1 µl
ddH <sub>2</sub> O	Ad 30 µl

Um Religationen zu umgehen, wurde dem Restriktionsansatz des Vektors im Fall von anschließender Ligation 1 µl Phosphatase (*Calf Intestine Phosphatase* CIP) zugegeben und nach Herstellerangaben inkubiert. Sowohl Insert als auch Vektor wurden vor Einsatz in die Ligation (II.4.5.2) via Macherey und Nagel-Kit aufgereinigt (II.4.4.2).

### **II.4.5.2    *Ligation***

Die Verknüpfung von DNS-Fragmenten mit kompatiblen Enden (II.4.5.1) geschah durch enzymatische Aktivität der T<sub>4</sub>-DNS Ligase von NEB. Dazu wurden Vektor und Insert in einem molaren Verhältnis von 1:3 bis 1:5 nach Herstellerangaben mit Puffer, Ligase und ddH<sub>2</sub>O vermischt und für mindestens 4 h bei 16°C inkubiert. Zur Selektion ligierter Plasmide wurden 5 µl des Ligationsansatzes in kompetente *Escherichi coli* –Zellen transformiert (II.4.7.2).

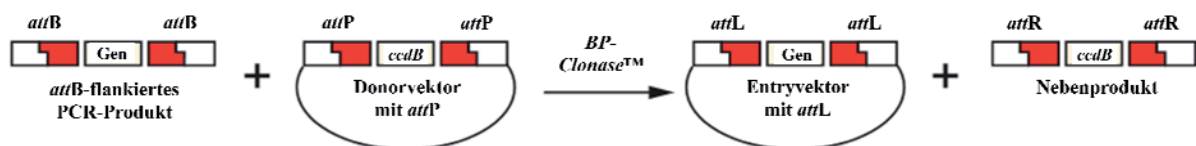
## **II.4.6      *In vitro -Rekombination***

Die im Folgenden beschriebene *in vitro* –Rekombination stellt eine Methode zur Subklonierung bzw. Fusionierung von DNS-Fragmenten dar, die ohne Restriktion und Ligation (II.4.5) auskommt, also unabhängig ist von einmalig vorkommenden Restriktionsschnittstellen innerhalb der beteiligten Fragmente. Sie basiert auf dem Einsatz isolierter Phagenenzyme, die die jeweiligen DNS-Fragmente über spezifische Rekombinationsseiten verbinden.

#### II.4.6.1 Gateway-Klonierung

Die Gateway-Technologie ist kommerziell von Invitrogen erhältlich. Sie basiert auf dem sequenzspezifischen Rekombinationssystem des Bakteriophagen Lambda, das es diesem ermöglicht ins *Escherichia coli* –Genom zu integrieren und dort vom lytischen in den lysogenen Zyklus zu wechseln. Die Rekombination erfolgt zwischen sogenannten *attachment-sites*, den Bindungsstellen der Rekombinationsproteine: *attB* auf dem bakteriellen Chromosom und *attP* auf dem Lambda-Chromosom. Sie wird von verschiedenen Enzymen katalysiert, die an die *attachment-sites* binden, die Zielsequenzen zusammenbringen, schneiden und nach Austausch die DNS kovalent verknüpfen. Die jeweiligen Enzyme variieren je nach lysogenem oder lytischem Zyklus. Die Rekombination der Bindungsstellen führt zu einem Austausch der innerhalb liegenden DNS-Sequenzen und zu modifizierten *attachment-sites* (Abbildung II.3 + Abbildung II.4). Die Gateway-Technologie nutzt dieses System für den erleichterten Transfer von heterologer DNS hinein in und zwischen Vektoren. Sie besteht grundsätzlich aus zwei Rekombinationsreaktionen, der BP- und der LR-Reaktion.

**BP-Reaktion:** Das Rekombinationsereignis wird katalysiert von der Phagen-Integrase und dem *E. coli* –Integrationsfaktor aus dem lysogenen Zyklus. Im Verlauf dieser Reaktion wird ein *attB*-flankiertes PCR-Produkt in einen *attP*-enthaltenden Donorvektor integriert. Die Bindungsstellen *attB* und *attP* vermischen sich dabei zu der Hybridsequenz *attL* (Abbildung II.3), welche in der anschließenden LR-Reaktion weiterrekombiniert (Abbildung II.4). Die Selektion erfolgt über das *ccdB*-Gen, welches für einen Gyrasehemmer codiert. Wurde dieses Gen im Donorvektor also nicht durch Rekombination mit dem zu klonierenden Gen ausgetauscht, sterben alle Zellen, die mit dem Vektor transformiert werden. Hat das Rekombinationsereignis allerdings stattgefunden, befindet sich das „Selbstmördergen“ *ccdB* ausschließlich im linearisierten Nebenprodukt, welches selbst im Falle einer Transformation in der Zelle abgedaut wird.



**Abbildung II.3:** BP-Reaktion der Gateway-Technologie, katalysiert vom BP-Clonase Mix. Das zu klonierende Gen wird durch Rekombination an den flankierenden *attB*-Sequenzen mit dem *ccdB*-Gen zwischen *attP*-Sequenzen im Donorvektor ausgetauscht. Es entstehen die neuen Hybridsequenzen *attL* im Entryklon, sowie *attR* im linearen Nebenprodukt

**LR-Reaktion:** Die zweite Reaktion des Gateway-Systems rekombiniert den in der BP-Reaktion generierten „Entry“-Vektor über die *attL*-Sequenzen mit den *attR*-Sequenzen eines beliebigen Expressionsvektors. Die Dabei entstehende Hybridsequenz aus *attL* und *attR* ist wieder die *attB*-site (Abbildung II.4). Katalysiert wird dieses Rekombinationsereignis von der Phagen-Integrase bzw. -Excisionase und dem bakteriellen Integrationsfaktor aus dem lytischen Zyklus.

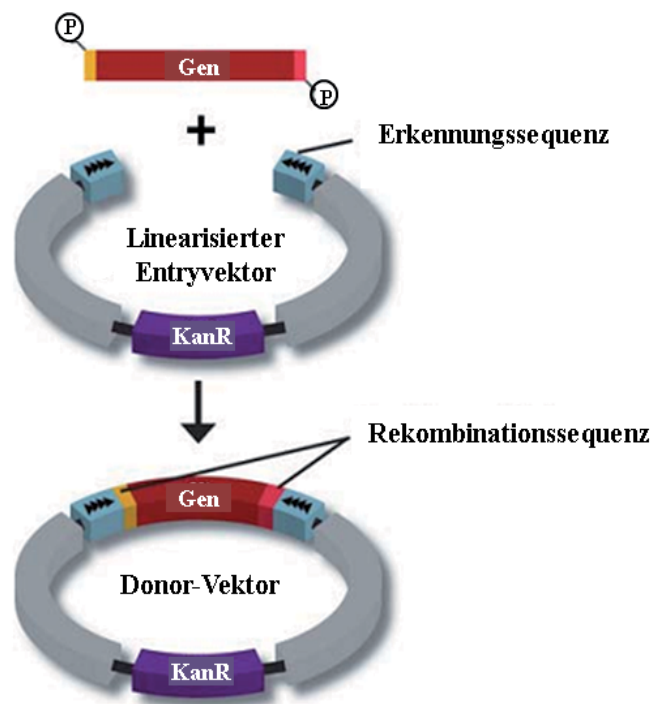


**Abbildung II.4:** LR-Reaktion der Gateway-Technologie, katalysiert vom LR-Clonase Mix. Das zu klonierende Gen im Entryvektor wird durch Rekombination an den flankierenden *attL*-Sequenzen mit dem *ccdB*-Gen zwischen *attR*-Sequenzen im Destinationsvektor ausgetauscht. Es entstehen die neuen Hybridsequenzen *attB* im Expressionsvektor, sowie *attP* im *ccdB*-enthaltenden Nebenprodukt.

Mittlerweile sind neben bakteriellen Anwendungen auch zahlreiche Destinationsvektoren für die Expression in Hefe kommerziell erhältlich. Sie unterscheiden sich in der Wahl des Selektionsmarkers, des Promotors und seiner Induzierbarkeit. Im Rahmen dieser Arbeit wurden ausschließlich *high-copy*-Plasmide mit konstitutiven Promotoren als Destinationsvektor eingesetzt (Tabelle II.2)

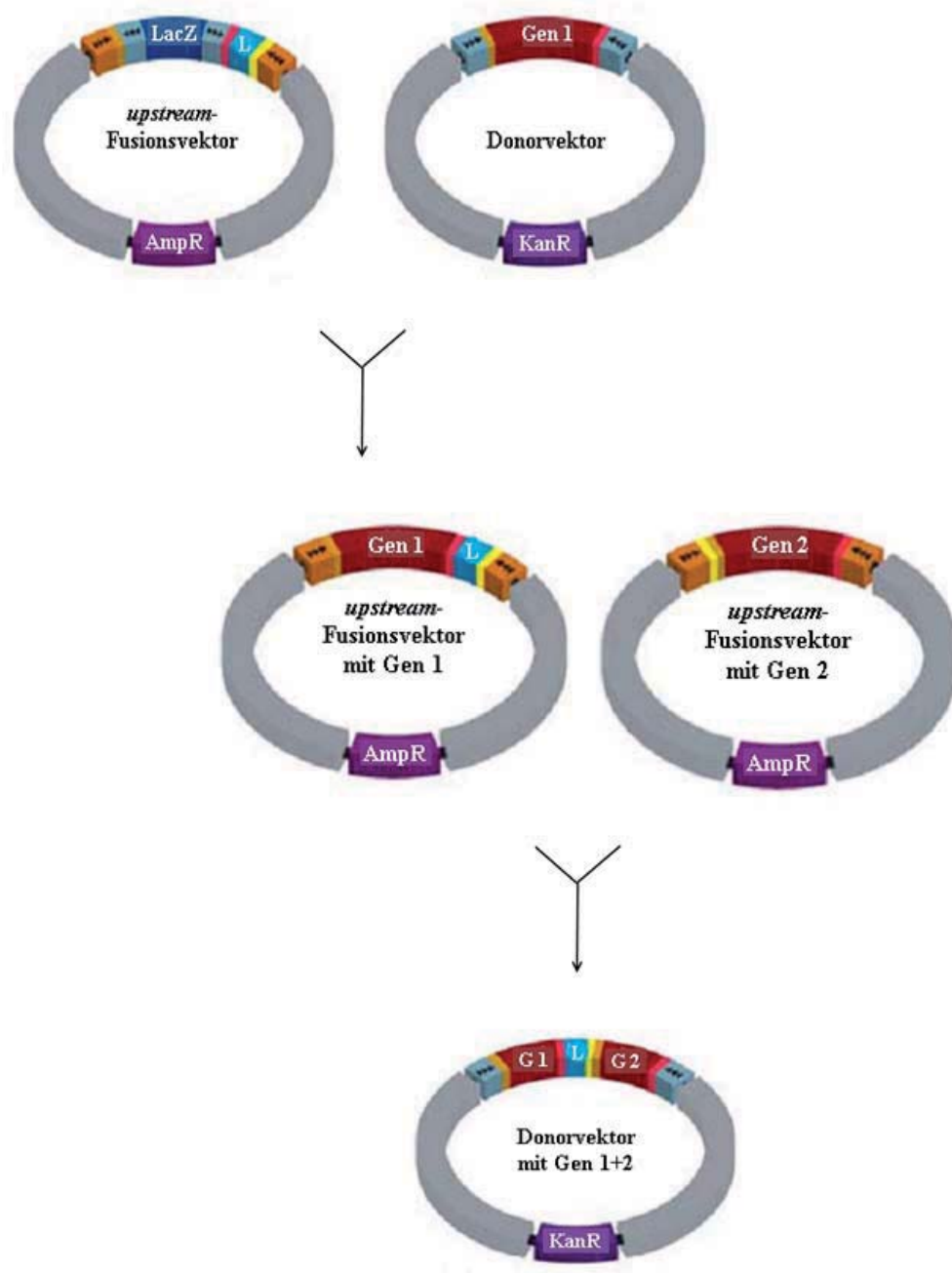
#### II.4.6.2 Stargate-Klonierung

Die Stargate-Technologie wird von IBA-BioTAGnology vertrieben. Sie basiert, wie auch die Gateway-Methode (II.4.6.1), auf dem sequenzspezifischen Rekombinationssystem des Bakteriophagen Lambda. Im Gegensatz zu den dort benötigten 25 bp-langen *attB*-Sequenzen, kommt dieses System allerdings mit Rekombinationssequenzen einer Länge von 4 Basenpaaren mit endständiger Phosphorylierung aus (Abbildung II.5).



**Abbildung II.5:** Generierung eines Donorvektors. Zur Re-kombination des DNS-Fragments in den linearisierten Entryvektor, müssen mittels PCR die 4 Basenpaar langen, phosphorylierten Rekombinationsseiten angehängt werden. **KanR:** Kanamycin-Resistenzgen.

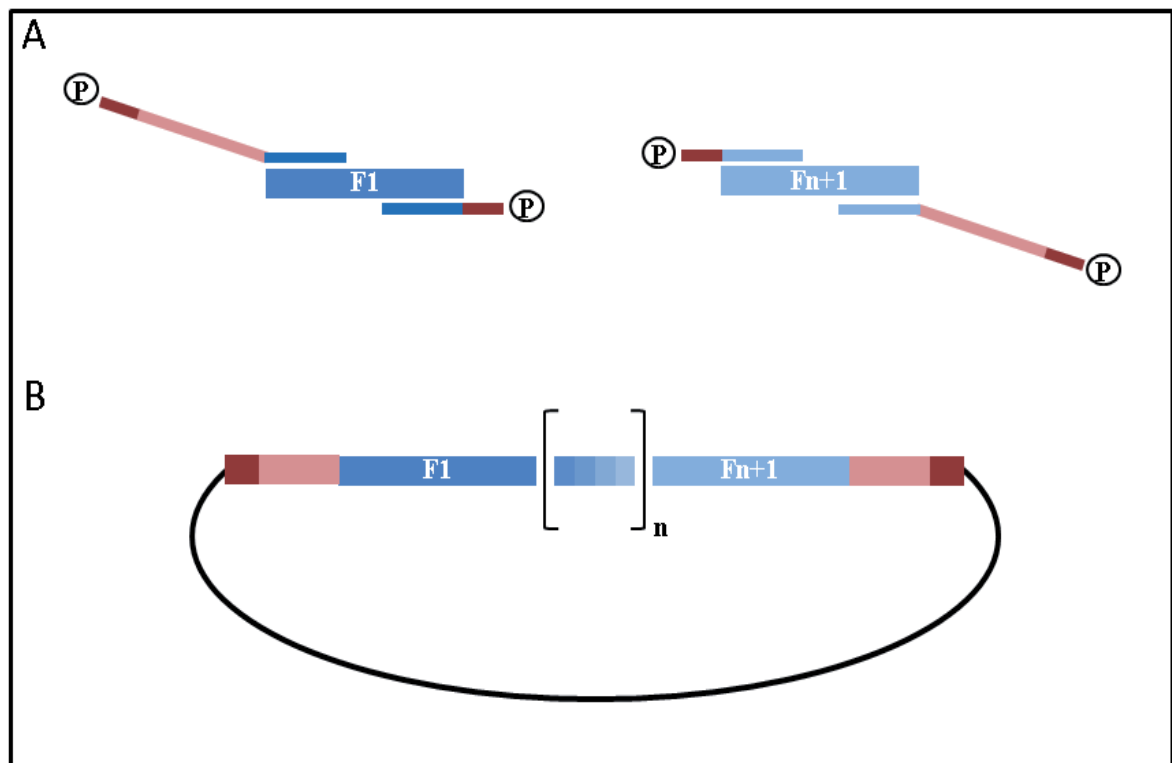
Ähnlich wie bei der Gateway-Technologie kann das zu klonierende Gen aus dem generierten Donorvektor auch beim Stargate-System in einen beliebigen Destinationsvektor mit der benötigten genetischen Umgebung rekombiniert werden. Als zusätzliche Strategie bietet diese Methode die Möglichkeit einer Fusionierung mehrerer Gene (Abbildung II.6). Dazu müssen diese vom Donorvektor in Fusionsvektoren rekombiniert werden. Je nach gewünschter Position innerhalb des geplanten Fusionskonstrukts, muss der generierte Donorvektor mit einem *upstream* - Fusionsvektor (für die 5'-Position) oder einem *downstream* - Fusionsvektor (für die spätere 3'-Position) reagieren. Diese rekombinieren anschließend zu einem Hybridplasmid mit den fusionierten Inserts. Verbunden sind die beiden DNS-Fragmente über einen im *upstream* - Fusionsvektor befindlichen Linker, der je nach Expression in Bakterien, Hefen oder tierischen Zellen einen Peptidlinker, eine ribosomale Bindestelle oder eine IRES-Sequenz beinhaltet.



**Abbildung II.6:** Fusionierung von DNS-Fragmenten mithilfe des Stargate-Systems. Nach Rekombination der zu verbindenden Fragmente in *upstream*- und *downstream*-Fusionsvektor, werden die beiden Inserts durch ein Rekombinationsereignis der beiden generierten Vektoren miteinander verbunden. Die verschiedenen Schritte sind über wechselnde Resistenzmarker selektierbar. **AmpR:** Ampicillin-Resistenzgen; **KanR:** Kanamycin-Resistenzgen; **L:** Linkerregion, über die die beiden Inserts verbunden werden; **LacZ:**  $\beta$ -Galaktosidasegen zur Blau-weiß-Selektion.



Das Stargate-System bietet leider keine große Auswahl an Destinationsvektoren für die rekombinante Expression in Hefen. Daher wurde im Rahmen dieser Arbeit nach der Fusionierung verschiedener DNS-Fragmente mithilfe der Stargate-Technologie eine Rekombination in Gateway-Destinationsvektoren angeschlossen. Die dafür benötigten Gateway-kompatiblen Rekombinationsseiten wurden den endständigen DNS-Fragmenten des späteren Fusionskonstrukts zusammen mit den Stargate-Rekombinationssequenzen durch PCR angehängt (Abbildung II.7).



**Abbildung II.7:** Kombination der Stargate- und Gateway-Technologie. **A:** Die einzelnen DNS-Fragmente werden separat in Stargate-Vektoren rekombiniert und miteinander fusioniert. Um eine anschließende Gateway-Klonierung zu ermöglichen, wird die dafür erforderliche Rekombinationssequenz bei den endständigen Fragmenten zwischen die Stargate- und die Fragmentsequenz eingefügt. Beim ersten Fragment (F1) erfolgt die Platzierung am 5'-Ende, beim letzten Fragment (Fn+1) am 3'-Ende. Dies geschieht bereits bei der Amplifikation der Fragmente für das erste Rekombinationsereignis, bei dem die phosphorylierte Stargate-Seite an das DNS-Fragment angehängt wird. **B:** Durch Fusionierung aller DNS-Fragmente über die Stargate-Technologie wird ein Vektor generiert, der über endständige Gateway-Seiten im Insert mit einem pDONR des Gateway- Systems rekombinierbar ist.

■: Stargate-Rekombinationsseite; ■: Gateway-Rekombinationssequenz; ■: DNS-Fragmente.

## **II.4.7      *Transfer von DNS***

### *II.4.7.1      Herstellung von dauerhaft kompetenten Escherichia coli –Zellen*

Zur Herstellung elektrokompeter *E. coli*-Stämme wurden diese im 500 ml-Maßstab mit einer Übernachtskultur beimpft und bis zu einer OD<sub>600</sub> von 0,5 inkubiert (II.3.1). Die pelletierten Zellen wurden 5 mal mit sterilem Eiswasser gewaschen, in 5 ml autoklaviertem 10% Glycerin aufgenommen und bis zum Gebrauch aliquotiert bei -80°C gelagert.

### *II.4.7.2      Transformation von E. coli-Zellen*

Die Transformation von *Escherichia coli*- Zellen erfolgte mit kommerziell bezogenen chemisch kompetenten Zellen nach Herstellerangaben oder durch den Einsatz selbst generierter elektrokompeter Zellen (II.4.7.1). Für die Elektrotransformation wurden die aliquotierten Zellen auf Eis aufgetaut und in einer sterilen, vorgekühlten Elektroküvette (0,2 cm) mit 50-200 ng Plasmid-DNS oder 5 µl eines Ligations- (II.4.5.2) oder Rekombinationsansatzes (II.4.6) vermischt. Nach der Elektroporation bei 2,5 kV für 5 msec wurden 700 µl SOC-Medium (II.2.2) zugegeben und der Ansatz zur Regeneration der Zellen und Ausbildung der benötigten Antibiotikumresistenz bei 37°C für 45 min inkubiert. Die Selektion erfolgte durch Ausstreichen von 100 µl der transformierten Zellen auf Agarplatten mit entsprechendem Antibiotikum.

### *II.4.7.3      Herstellung kompetenter Hefezellen*

Die Herstellung kompetenter Hefezellen erfolgte anhand der LiAc-Methode nach AMBERG *et al.* Dazu wurden 50 ml YPD-Medium (II.2.3) mit einer Übernachtskultur angeimpft und bei 28°C bis zu einer OD<sub>600</sub> von 1 kultiviert. Im Anschluss an einmaliges Waschen der pelletierten Zellen mit sterilem ddH<sub>2</sub>O und Resuspendieren in 100 mM LiAc wurden die Zellen in ein 1,5 ml Eppendorfgefäß überführt und für 10 sec bei 16100 xg zentrifugiert. Das Pellet wurde in 500 µl 100 mM LiAc gelöst und in 50 µl-Aliquots aufgeteilt. Die so generierten kompetenten Hefezellen sind ausschließlich für den sofortigen Einsatz in eine Transformation (II.4.7.4) gedacht und können nicht gelagert werden.

#### II.4.7.4 Transformation von Hefezellen

Für die Transformation von DNS in Hefezellen wurden ausschließlich frische kompetente Zellen verwendet (II.4.7.3). Dazu wurde ein 50 µl-Aliquot pro Transformationsansatz durch 10 sekündiges Zentrifugieren bei 16100 *xg* pelletiert und der Überstand verworfen. Folgende Reagenzien wurden anschließend in angegebener Reihenfolge zugegeben:

PEG 50%	240 µl
LiAc 1M	35 µl
ssDNS 2 mg/ml (denaturiert)	25 µl
DNS	500-1000 ng
ddH <sub>2</sub> O	ad 350 µl

Nach Mischen des Ansatzes (Vortexer), Inkubation bei 30°C für 30 min und einem Hitzeschock für 25 min bei 42°C, wurden die transformierten Zellen für 15 sec bei 8000 *xg* gesammelt, in 100 µl steriles ddH<sub>2</sub>O aufgenommen und auf Selektionsplatten (II.2.4) ausgestrichen. Als Negativkontrolle diente ein Ansatz ohne DNS-Zugabe. Nur bei fehlenden Kolonien auf der negativen Selektionsplatte wurden Kolonien der Transformationsplatte für weitere Untersuchungen verwendet.

#### II.4.7.5 Bestimmung der Transformationseffizienz

Da bei kommerziell erhältlichen kompetenten Zellen die Transformationseffizienz vom Hersteller bestimmt wird, wurde die Effizienz von *E. coli* –Zellen ausschließlich bei Gebrauch der selbst generierten kompetenten Zellen bestimmt. Dies erfolgte anhand folgender Formel:

$$\frac{\text{Anzahl Kolonien auf Selektionsplatte}}{\mu\text{g an ausplattierter DNS}} \geq 10^6$$

Ergab die berechnete Effizienz einen Wert unter 10<sup>6</sup> Zellen / µg DNS, kamen die getesteten Zellen nicht zum Einsatz.

## II.5 Biochemische Methoden

### II.5.1 Isolierung von Gesamtprotein

#### II.5.1.1 Herstellung von Rohextrakt

Rohextrakte wurden ausschließlich von Hefestämmen zur anschließenden Proteinanalyse hergestellt. Dazu wurden die Zellen im 15 ml Maßstab angezogen, geerntet (II.3.2) und in 500 µl Aufschlusspuffer resuspendiert. Nach Zugabe von 0,2 g Glasperlen (0,2-0,5 mm Durchmesser, Roth) und kurzem Vortexen zur Abtrennung von äußeren Membranproteinen folgte ein Zentrifugationsschritt für 30 sec bei 16100  $xg$  zum Austausch des Puffers. Der Aufschluss der Zellen erfolgte anhand einstündigen Vortexens bei 4°C. Die Glasperlen und Zelltrümmer wurden durch einen Zentrifugationsschritt (1 min, 16100  $xg$ ) abgetrennt und der Überstand auf seinen Proteingehalt untersucht.

<b>Aufschlusspuffer:</b>	NaPHO <sub>4</sub>	25mM
	DTT	2mM
	pH7,5	

#### II.5.1.2 Bestimmung der Proteinkonzentration

Zur Quantifizierung isolierter Proteine wurde der Bradfordtest angewandt. Dazu wurden die Proben 1:50 mit ddH<sub>2</sub>O verdünnt, 10 µl davon mit 500 µl Bradfordreagenz (Roth, 1:5 verdünnt und filtriert) vermischt und nach einer Inkubationszeit von mindestens 8 min spektrophotometrisch die OD<sub>595</sub> bestimmt. Alle Proben wurden im Dreifachansatz bestimmt. Als Nullprobe dienten 10 µl ddH<sub>2</sub>O, die ebenfalls mit 500 µl Bradfordreagenz vermischt und für 10 min inkubiert wurden. Um die bestimmten OD<sub>595</sub>-Werte in Proteinmengen umrechnen zu können, wurde eine Standardgerade erstellt. Dazu wurde eine serielle Verdünnung zwischen 500 und 50 µg BSA / ml angefertigt und mittels Bradfordtest der jeweilige OD<sub>595</sub>-Wert detektiert. Die Messwerte wurden gegen die eingesetzte BSA-Konzentration aufgetragen und mit der generierten Geradengleichung die detektierten OD<sub>595</sub>-Werte der Proteinproben in die entsprechende Proteinkonzentration umgerechnet.

#### II.5.1.3 Proteinfällung

Für die effektive Auftrennung von Gesamtprotein aus Hefezellen durch SDS-PAGE (II.5.2) wurde diese durch eine TCA-Fällung entsalzt und konzentriert. Dafür wurden 150µg Gesamtprotein (0) 5:1 mit TCA (50% w/v) vermischt, die Proteine durch einminütiges Zentrifugieren bei 16100  $xg$

pelletiert und in 32µl SDS-Laufpuffer (II.5.2) resuspendiert. Die Denaturierung erfolgte durch Zugabe von 8µl vierfachkonzentrierten SDS-Probenpuffers (II.5.2) und Erhitzen auf 95°C.

### II.5.2 SDS-Polyacrylamidgelelektrophorese

Um aufgereinigte Proteinfractionen (II.5.1.3) zu analysieren, wurden diese gelelektrophoretisch ihrer Größe nach aufgetrennt und anschließend durch Anfärben mit Coomassie visualisiert. Als Trennmethode diente dabei ein diskontinuierliches Gel auf Polyacrylamidbasis bestehend aus einem 4%-igen (w/v) PAA-Sammelgel und einem 10%-igen (w/v)-PAA-Laufgel.

<b>Sammelgel:</b>	Tris-HCl 1M pH6,8	625 µl
	Polyacrylamid 30% (v/v)	830 µl
	SDS 10%	50 µl
	TEMED	5 µl
	APS 20% (w/v)	15µl
	ddH <sub>2</sub> O	3625 µl
<b>Trenngel:</b>	Tris-HCl 1M pH8,8	3750 µl
	Polyacrylamid 30% (v/v)	3300 µl
	SDS 10%	100 µl
	TEMED	10 µl
	APS 20% (w/v)	30 µl
	ddH <sub>2</sub> O	2785 µl

Analysiert wurde jeweils eine Menge von 40µg Gesamtprotein (II.5.1.3), die zur Auflösung von Sekundär-, Tertiär- und Quartärstruktur vor Beladen des Gels in SDS-Probenpuffer bei 95°C denaturiert wurden. Das darin enthaltene anionische Tensid SDS überdeckt zudem die Eigenladung der Proteine, sodass sie sich im angelegten elektrischen Feld ausschließlich in Abhängigkeit ihrer Größe bewegen.

<b>SDS-Probenpuffer 10x:</b>	Tris-HCl pH6,8	62,5 mM
	Glyzerin	30% (w/v)
	SDS	4% (w/v)
	β-Mercaptoethanol	10% (w/v)
	Bromphenolblau	0,05% (w/v)
<b>SDS-Laufpuffer:</b>	Tris	125 mM
	Glyzerin	960 mM
	SDS	0,5% (w/v)
	pH8,3	

Als Größenstandard dienten 5 µl einer Proteinleiter (Promega). Nach gelelektrophoretischer Auftrennung in einer Mini-Protean Laufkammer (Biorad) bei 180 V für 45 min wurde das Sammelgel entfernt und die Proteine im Trenngel nach WONG *et al*, 2000 visualisiert. Diese Methode besteht aus 4 Färbeschritten, die jeweils in der Mikrowelle durchgeführt wurden und sowohl die Coomassie-Färbung als auch die Entfärbung des Gels beinhalten. Das Gel wurde dazu nacheinander in Lösung A-D gegeben, bei 800 Watt bis zum Kochen der jeweiligen Flüssigkeit erhitzt und vor Inkubation in der nächsten Lösung mit Wasser gewaschen. Im Anschluss an die Behandlung mit Lösung D wurde das Gel nicht gewaschen, sondern ein Stück Papier zugegeben, um Restfärbung abzufangen.

<b>Lösung A:</b>	Coomassie	0,05% (w/v)
	Isopropanol	25% (v/v)
	Essigsäure	10% (v/v)
<b>Lösung B:</b>	Coomassie	0,005% (w/v)
	Isopropanol	10% (v/v)
	Essigsäure	10% (v/v)
<b>Lösung C:</b>	Coomassie	0,002% (w/v)
	Isopropanol	10% (v/v)
<b>Lösung D:</b>	Isopropanol	10% (w/v)

Zur Datenaufbewahrung wurden alle entfärbten Gele eingescannt.

### II.5.3 *Westernblot*

Zur Immunoblotanalyse wurden die gelelektrophoretisch aufgetrennten Proteine (II.5.2) nicht angefärbt, sondern auf eine Nitrocellulosemembran (0,45 µm, Millipore) transferiert. Durch Anlegen eines zum Gel senkrecht gerichteten elektrischen Feldes wandern die Proteine auf die Membran und haften dort an der hydrophoben Oberfläche. Der Transfer erfolgte bei 120 V für 90 min in einer eisgekühlten „Tank-Blot“-Apparatur von BioRad.

<b>Transferpuffer:</b>	Glyzin	1,44% (w/v)
	Tris	0,3% (w/v)
	Methanol	20% (w/v)

Die anschließende zweistufige Detektion der Zielproteine basierte auf dem Antigen-Antikörperbindungsprinzip, wobei der eingesetzte Zweitantikörper mit Alkalischer Phosphatase (AP) markiert war. Bei spezifischer Bindung dieses Antikörpers kann das Zielprotein dabei durch die Umsetzung des zugegebenen artifiziellen Substrats NBT/BCIP zu einem blauen Indigofarbstoff lokalisiert und anhand des mitgeführten Standards seine Größe bestimmt werden. Das NBT/BCIP (Nitro-Blau Tetrazolium Chlorid NBT/5-Brom-4-Chloro-3-Indoxylphosphat BCIP) wurde von Roth bezogen, 2:1 vermischt und 20%ig (w/v) in Formamid bei -20°C gelagert. Die genaue Vorgehensweise bei Durchführung des Westernblots war wie folgt:

Schritt	Funktion	Reagenz	Zeitintervall
1	Blockieren der freien Bindungsstellen auf der Membran	5% Milchpulver in PBS-T	90 min
2	Spezifische Bindung des Erstantikörpers an Zielprotein	Rabbit anti His-Antikörper	60 min
3	Abwaschen unspezifisch gebundener Antikörper (AK)	PBS-T	3 x 15 min
4	Spezifische Bindung des Zweitantikörpers an Erst-AK	Goat anti rabbit, AP-markiert	60 min
5	Abwaschen unspezifisch gebundener Zweit-AK	PBS-T	3 x 15 min
6	Äquilibrierung der Membran für Nachweisreaktion	AP-Puffer	20 min
7	Nachweisreaktion durch Alkalische Phosphatase	NBT/BCIP (Substrat) 1:100 in AP-Puffer	bis Detektion von Banden
8	Abstoppen der Farbreaktion	Wasser	5 min

Die für die Durchführung eines Westernblots benötigten Lösungen setzten sich folgendermaßen zusammen:

<b>PBS 10x:</b>	NaCl	1370 mM
	KCl	27 mM
	Na <sub>2</sub> HPO <sub>4</sub>	81 mM
	KH <sub>2</sub> PO <sub>4</sub>	15 mM

<b>PBS-T:</b>	Tween 20	0,05% (v/v)
	PBS 1x	

<b>AP-Puffer:</b>	Tris-HCl	100 mM
	NaCl	100 mM
	MgCl <sub>2</sub>	5 mM
	pH9,6	

### II.5.4 Northernblot

Um die Expression rekombinanter Gene in generierten Hefestämmen zu analysieren, wurde ein Northernblot durchgeführt und die rekombinante mRNA durch Hybridisierung mit komplementären Gensonden markiert. Als Negativkontrolle diente der Hefestamm-Wildtyp, an dessen RNS die Sonden nicht binden sollten.

#### II.5.4.1 Gelelektrophoretische Auftrennung der RNS

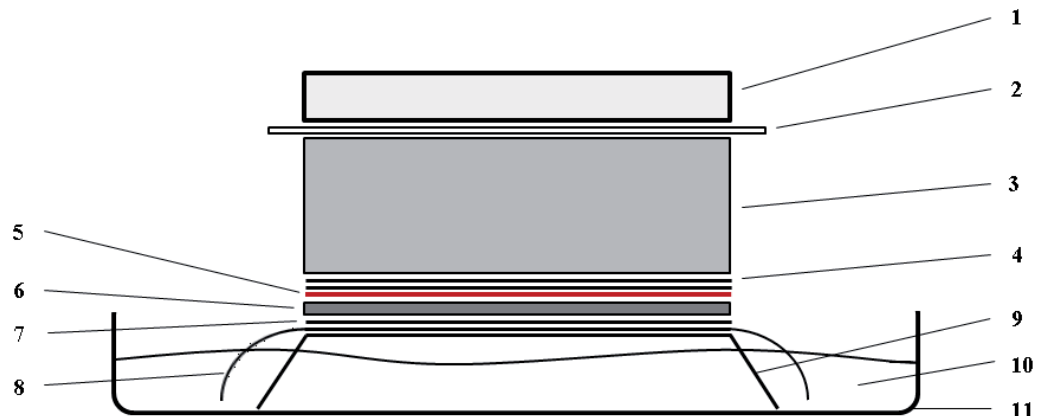
Zehn µg isolierter Gesamt-RNS (II.4.1.3) wurden mit gleichem Volumen an Probenpuffer für 10 min bei 80°C denaturiert und der Ansatz in einem Formaldehyd-haltigen 1,5% (w/v) Agarosegel bei 80 Volt für 120 Minuten gelelektrophoretisch aufgetrennt. Mithilfe des Transilluminators wurde die Ethidiumbromid-markierte RNS der einzelnen Proben auf vergleichbare Konzentrationen und gleichmäßige Auftrennung im Gel überprüft.

<b>Probenpuffer:</b>	Laufpuffer	1,5 ml
	Formaldehyd	2,4 ml
	Glyzerin	3,0 ml
	Formamid	8,1 ml
	Bromphenolblau	Spatelspitze
	Ethidiumbromid	17 µl
<b>Laufpuffer 10x :</b>	MOPS	41,8 g
	DEPC-Wasser	700 ml
	pH7,0 mit 2N NaOH	
	1M Na-Acetat	20 ml
	EDTA 500mM (pH8,0)	20 ml
	DEPC-Wasser	ad 1000 ml
<b>DEPC-Wasser:</b>	DEPC	0,1% (v/v)
	Rühren über Nacht	
	Autoklavieren	

#### II.5.4.2 Transfer der RNS auf Nylonmembran

Um eine spezifische Markierung der rekombinanten mRNA zu ermöglichen, wurden die aufgetrennten Nukleinsäuren über Nacht auf eine Nylonmembran (Hydro+, GE Healthcare) transferiert. Die einzelnen Bestandteile der Blotapparatur wurden wie in Abbildung II.8 dargestellt aufeinander geschichtet und über Nacht bei Raumtemperatur inkubiert.





**Abbildung II.8:** Aufbau eines Northernblots. **1:** Gewicht; **2:** Glasplatte zur gleichmäßigen Gewichtsverteilung; **3:** Papierstapel für Kapillarsog; **4:** Zwei Lagen Whatmanpapier; **5:** Hydrobond+ Nylonmembran; **6:** Gel mit gelelektrophoretisch aufgetrennter RNS; **7:** Eine Lage Whatmanpapier; **8:** eine Lage Whatmanpapier mit überlappenden, in den Transferpuffer reichenden Seiten; **9:** umgedrehter Gelträger; **10:** Transferpuffer; **11:** Glasschale

Nach Transfer auf die Membran in 10x SSC wurde die RNS für 30min bei 80°C hitzefixiert und über Visualisierung der Ethidiumbromidmarkierung im Transilluminator auf quantitativen Transfer überprüft.

<b>SSC 20x:</b>	NaCl	175,3 g
	Na-Zitrat	88,2 g
	DEPC-Wasser	800 ml
	pH7,0	
	DEPC-Wasser	ad 1000 ml

#### II.5.4.3 Hybridisierung

Während der 2-3 stündigen Prähybridisierung der Membran bei 42°C wurden die zur Hybridisierung an die rekombinante RNS benötigten Oligos radioaktiv markiert. Dies erfolgte anhand der T<sub>4</sub>-Polynukleotidkinase PNK (NEB), die den Transfer eines Phosphats von der  $\gamma$ -Position eines  $\gamma$ -<sup>32</sup>P-ATP-Moleküls (Hartmann Analytik) auf den 5'-Hydroxyterminus eines Polynukleotids katalysiert.

<b>Prähybridisierungspuffer:</b>	ULTRAhyb (Applied bisoystems)	10 ml
	SSC 6x	5 ml
	SDS 7%	5 ml
	65°C	

<b>Markierungsansatz:</b>	PNK-Puffer	5 µl
	PNK	20 U
	Oligo	20 pmol
	$\gamma$ - <sup>32</sup> P-ATP	1,85 MBq
	ddH <sub>2</sub> O	ad 50µl

Der Ansatz wurde für 30-45 min bei 37°C inkubiert und die radioaktiv markierten Oligos mithilfe des *Nucleotide removal kits* von Qiagen von überschüssigen  $\gamma$ -<sup>32</sup>P-ATP getrennt. Ab Zugabe der Radioaktivität fanden aus sicherheitstechnischen Gründen alle anschließenden Arbeitsschritte im Isotopenlabor des Fraunhofer IME statt. Die generierten Sonden wurden dem Prähybridisierungsansatz zugegeben und zur Bindung an die Membran für 24 h in einem Rotationsofen bei 42°C inkubiert. Um unspezifisch hybridisierte Oligos von der Membran zu trennen, wurde anschließend zweimal für 20 min mit 2x SSC / 0,1% SDS und einmal mit 0,1x SSC / 0,1% SDS gewaschen.

#### II.5.4.4 Visualisierung

Der autoradiographische Nachweis gebundener Sonden erfolgte durch Auflegen der getrockneten Membran auf eine Phosphorplatte oder einen angeblitzten Röntgenfilm (Kodak). Die Beschichtung des Films mit fotografischer Emulsion auf Silberbromidbasis führte an Stellen einer Belichtung zur Reduktion der Silberionen zu metallischen Silberatomen. Um solche Reaktionen aufgrund nicht-radioaktiver Strahlung auszuschließen, wurden alle Arbeitsschritte zur Visualisierung bei Rotlicht durchgeführt. Nach Lagerung in einer licht- und röntgenstrahlungsdichten Kassette bei -80°C für 48 h, wurden die Silberkristalle durch Inkubation in GBX-Entwicklerlösung (Kodak) vergrößert. Die abschließende Fixierung in einer Natriumthiosulfat-haltigen Lösung (Kodak) resultierte in der Überführung verbliebenen Silberbromids in eine leicht lösliche Form und diente der dauerhaften Lichtbeständigkeit des Bilds. Alle entwickelten Röntgenfilme wurden zur weiteren Datenspeicherung eingescannt. Bei Einsatz der Phosphorplatte wurde diese mit der aufgelegten Membran ebenfalls für 48h in einer licht- und röntgenstrahlungsdichten Kassette bei -80°C inkubiert. Die angeregten Kristalle der Phosphorplatte wurden anschließend mithilfe des Phosphoimagers (Biorad) gescannt und die Emissionsintensität detektiert. Diese Methode bietet den Vorteil fehlender Überexpressionsrisiken, zeigt allerdings auch eine geringere Auflösung.

### II.5.5 *Massenspektrometrische Analyse*

Zur qualitativen und quantitativen Analyse der Hefestämme auf ihre Produktion an Taxadien, GGOH und Squalen wurden die zu untersuchenden Zellen und das filtrierte RP C18-Material nach der Sublimationstrocknung (II.3.2) auf eine Chromatographiesäule mit Glasfritte gegeben. Die Extraktion der gebildeten Produkte erfolgte durch Zugabe von 10 ml Cyclohexan. Um eine spätere Quantifizierung der Extrakte zu ermöglichen, enthielt das Lösungsmittel 20 mg/l 4,4-di-*tert*-butylbiphenyl als internen Standard. Zur vollständigeren Extraktion wurde das Eluat erneut auf die Säule gegeben, nach einer kurzen Inkubation gesammelt und 1 ml zur GC/MS-Analyse in eine Gewindeflasche mit Schraubkappe (VWR) überführt.

Die Extraktion aus im 30l-Fermenter gewonnener Biomasse (II.3.3) erfolgte mithilfe einer Soxhlet-Apparatur. Die Zellen wurden dafür in eine Cellulosehülse überführt und die gebildeten Produkte durch eine 24 stündige Extraktion mit 1l n-Pentan herausgelöst. Das Lösungsmittel wurde durch Begasung mit Druckluft eingedampft und die so konzentrierten Substanzen anhand GC/MS untersucht.

#### II.5.5.1 *GC/MS-Analyse*

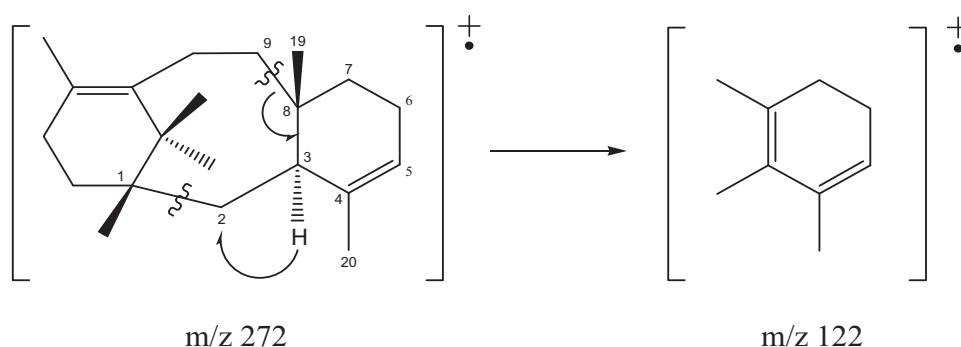
GC/MS-Analysen wurden an einem Shimadzu QP2010S-Gerät mit slit/slitless Injektor durchgeführt. Die gaschromatographische Auftrennung der Cyclohexanextrakte erfolgte dabei über eine 30 m lange, unpolare Rxi™-5ms-Säule der Firma Restek (Bad Homburg) mit einem Innendurchmesser von 0,25 mm. Als mobile Phase diente das Trägergas Stickstoff. Folgende Trennmethode kam dabei zum Einsatz:

Schritt	Temperatur	Zeitintervall
Injektion	250°C	
Gradient	200-270°C	4°C / min (22,5 min)
Finale Haltephase	270°C	2 min

Zur Identifizierung der einzelnen Komponenten des Stoffgemisches wurde an die gaschromatographische Auftrennung eine massenspektrometrische Messung gekoppelt. Die Ionisierung erfolgte dabei durch Elektronenstoß (EI) bei 1 keV, die Detektion über einen Quadropol-Analysator. Anhand der entstandenen Bruchstücke konnte mittels Literaturdaten auf die Strukturformel der untersuchten Substanzen rückgeschlossen werden. Als zusätzliche Messmethode

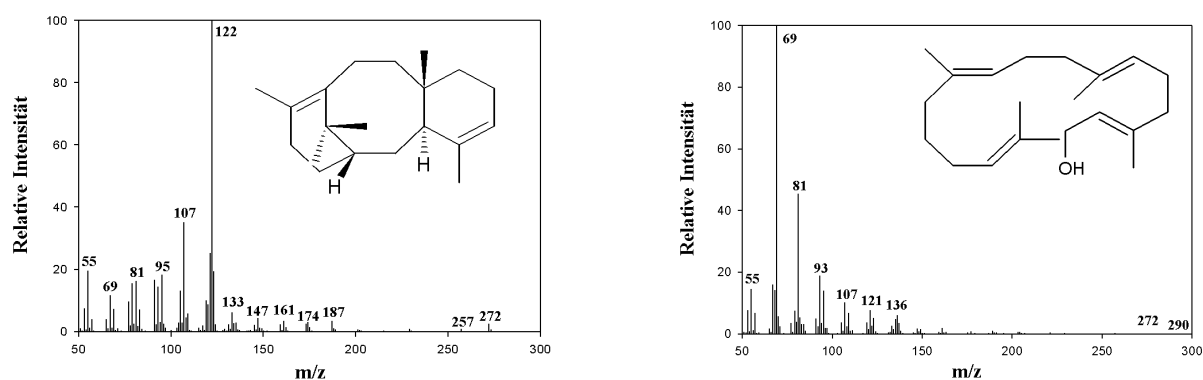
wurde neben dem *Scan*-Modus das SIM (*selective ion monitoring*) für charakteristische Massen der gesuchten Substanzen angewendet.

Die Identifizierung von Taxadien erfolgte dabei anhand des bekannten Massenspektrums 122(100); 107(32%), 257(1%) mit einem Molekülion von 272 (Koepp et al, 1995, Lin et al, 1996). Die Quantifizierung wurde mithilfe fermentativ hergestellter Substanz ermittelt und basierte auf dem Massenfragment  $m/z$  122. Dieses entsteht durch eine doppelte Bindungsspaltung (C9-C9), bzw. C1-C2) bei gleichzeitiger H-Umlagerung von C3 nach C2, was zu der Bildung einer Doppelbindung zwischen C8 und C3 führt (Abbildung II.9).



**Abbildung II.9: Entstehung des charakteristischen Massenfragments  $m/z$  122 bei der Fragmentierung von Taxadien.**<sup>[82]</sup> Durch Bindungsspaltungen zwischen C8 und C9, bzw. C1 und C2 und gleichzeitiger H-Umlagerung von C3 nach C2 entsteht eine Doppelbindung zwischen C8 und C3.

Sowohl die Identifizierung als auch die Quantifizierung von GGOH erfolgte durch den Vergleich mit der kommerziell erhältlichen Reinsubstanz (Roth). Diese wurde weiterhin für die Erstellung einer Standardgeraden zur Quantifizierung verwendet. Abbildung II.10 zeigt die Massenspektren von Taxadien und GGOH.



**Abbildung II.10: Massenspektren von Taxadien (links) und Geranylgeraniol (rechts).** Aufgetragen sind die relativen Intensitäten [%] in Abhängigkeit der Massen. Taxadien:  $m/z$  122(100%), 107(32%), 257(1%) mit einem Molekülion von  $m/z$  272. GGOH:  $m/z$  69(100%), 93(20%), 55(12%), 121(10%), 272(0,1%) mit einem Molekülion von  $m/z$  290.

## II.6 Verwendete Programme

Für die Auswertung von Ergebnissen wurden die in Tabelle II.6 aufgelisteten Programme verwendet.

**Tabelle II.6: Für die Auswertung verwendete Computerprogramme**

Bezeichnung	Verwendungszweck
Clone Manager Suite 8	Erstellung von Vektorkarten; virtuelle Restriktion & Ligation
Graphical codon usage analyser 2.0	Analyse des Codongebrauchs einer Gensequenz und Vergleich mit desjenigen des Expressionssystems
DNASTAR Lasergene Seqman	Auswertung von Sequenzierungsergebnissen; Abgleich verschiedener DNS-Fragmente auf homologe Bereiche
OriginPro 8G	Erstellung von Balkendiagrammen

## II.7 Chemikalien und Verbrauchsmaterial

### II.7.1 Chemikalien und Enzyme

Die im Rahmen dieser Arbeit verwendeten Chemikalien und Enzyme wurden von folgenden Firmen bezogen:

BioRad (München), IBABioTAGnology (Göttingen), Hartmann Analytik (Braunschweig), Invitrogen (Leek, Niederlande), Kodak (Stuttgart), New England Biolabs (Frankfurt am Main), Peqlab (Erlangen), Promega (Mannheim), Roth (Karlsruhe), Qiagen (Hilden), Serva (Heidelberg), Sigma (Deisenhofen), Stratagene (La Jolla CA, USA)

### II.7.2 Verbrauchsmaterialien

Die Verbrauchsmaterialien stammten von den Firmen:

Applied biosystems (Darmstadt), Eppendorf (Hamburg), GE Healthcare (Freiburg), Greiner (Solingen), Kodak (Stuttgart), Milian (Genf, Schweiz), Millipore (Eschborn), Roth (Karlsruhe), Sarstedt (Nümbrecht), Serva (Heidelberg), VWR (Darmstadt), Whatman (Bender & Hobein, Bruchsal).

### III Resultate

#### III.1 Prinzipielle Überlegungen

Die erfolgreiche Produktion terpenoider Verbindungen basiert nicht nur auf der heterologen Expression der entsprechenden pflanzlichen Biosynthesegene. Wie bereits angedeutet, ist sie ebenso abhängig von der Wahl des Wirtsorganismus, sowie der Optimierung seines Stoffwechsels zur Bereitstellung der benötigten Vorstufen. Weiterhin wird sie beeinflusst von der Art der Expression, Integration und Kopienzahl der jeweiligen Gene. Im Folgenden sollen daher grundlegende Überlegungen beschrieben werden, die vor Durchführung der Arbeiten angestellt wurden.

##### III.1.1 Expressionssystem und Wahl des Stammes

Wohl eine der wichtigsten Überlegungen zur Durchführung dieser Promotionsarbeit war die Wahl des Organismus, mit dessen Hilfe die rekombinante Produktion der taxoiden Verbindungen durchgeführt werden sollte. Wie bei pflanzlichen Sekundärstoffen üblich, basiert auch die Biosynthese von Taxol auf Cytochrom-P450-vermittelten Modifikationsreaktionen des Taxadien Grundgerüsts. So werden ungefähr die Hälfte aller enzymatischen Schritte des Synthesewegs durch diese Enzymklasse katalysiert.<sup>[93-94]</sup> Für die Selektion des im Rahmen dieser Arbeit verwendeten Expressionssystems ist die Fähigkeit, in die Taxolbiosynthese involvierte Cytochrom-P450-abhängige Monooxygenasen funktionell zu exprimieren, daher unumgänglich.

Eukaryotische Cytochrom-P450 Monooxygenasen sind durch eine N-terminale Transmembranhelix in der ER-Membran fixiert<sup>[97]</sup> und mit einer NADPH-Reduktase für den benötigten Elektronentransfer assoziiert.<sup>[96]</sup> Ausnahmen zu dieser üblichen Positionierung bilden vier in *Arabidopsis thaliana* nachgewiesene Enzyme, die in der inneren oder äußeren Chloroplastenmembran verankert sind.<sup>[162]</sup> Im Gegensatz zu ihren eukaryotischen Pendants liegen bakterielle P450 in löslicher Form vor.<sup>[163]</sup> Für die heterologe Expression pflanzlicher P450 in Bakterien müssen diese daher N-terminal modifiziert werden. Ungeachtet dessen wurden grundsätzliche Probleme wie Abweichungen in Faltung und Translation, eine benötigte korrekte Insertion in die Membran, fehlende posttranslationale Modifizierungen und Assoziierung mit Proteinpartnern bereits beobachtet.<sup>[146]</sup> Das für eine rekombinante Expression wohl meist eingesetzte Bakterium *Escherichia coli* weist selbst keine Cytochrom-P450-Enzyme auf. Und obwohl dieser Modellorganismus die löslichen Flavoproteine Flavodoxin und NADPH-Flavodoxin-Reduktase synthetisiert, stellten letztere bisher nicht für alle untersuchten heterologen P450 die

benötigten Reduktionsequivalente bereit.<sup>[164]</sup> Die wenigen erfolgreich in *Escherichia coli* exprimierten P450 führten daher ausschließlich bei Zugabe einer exogenen NADPH:Cytochrom-P450-Reduktase<sup>[164-168]</sup> zu einem detektierbaren Produkt, oder aber bei Konstruktion eines Fusionsproteins aus P450 und der entsprechenden NADPH:Cytochrom-P450-Reduktase.<sup>[144][145]</sup> Eine effiziente Terpenoxidation *in vivo* konnte trotz funktionell exprimierter pflanzlicher P450-Gene allerdings nicht gezeigt werden.<sup>[146]</sup> Während der Durchführung dieser Arbeit wurde dennoch die erste *in vivo*-Produktion von funktionalisierten Terpenen in *E. coli* gezeigt. Realisiert wurde dies durch die Expression einer Codon-optimierten Hydroxylase, eines *N*-terminal angefügten Membranankers, der gleichzeitigen Expression einer NADPH-Cytochrom-P450-Reduktase und Optimierung des *E. coli*-eigenen Primärstoffwechsels für ausreichende Mengen an Vorstufen. Die Produktionsmengen variierten dabei abhängig von der koexprimierten Reduktase, des gewählten *E. coli*-Stamms und des Vektorsystems.<sup>[169]</sup> Zum Zeitpunkt der Wahl des Expressionssystems allerdings zeigte *E. coli* im direkten Vergleich mit *S. cerevisiae* bei rekombinanter Expression pflanzlicher P450 sowohl geringere spezifische Aktivitäten, als auch geringe Hydroxylierungsraten.<sup>[167]</sup> Von dem Einsatz eines prokaryotischen Systems wurde daher abgesehen und stattdessen *S. cerevisiae* verwendet, das bis dato erfolgreichste System für die rekombinante Produktion von terpenoiden Verbindungen wie Pregnenolon,<sup>[170]</sup>  $\beta$ -Carotin und Leukopen<sup>[124]</sup> bzw. Hydrocortison.<sup>[171]</sup> *S. cerevisiae* gewährleistet nicht nur die Vorteile eines einzelligen Organismus bei gleichzeitiger Proteinprozessierung eines Eukaryoten, sondern wurde des Weiteren von der *American Food and Drug Administration* als GRAS-Organismus eingestuft (**G**enerally **R**egarded **A**s **S**afe), was die Produktion pharmazeutischer Verbindungen mit der Bäckerhefe erleichtert. Aufgrund seiner Bedeutung auf dem Gebiet der Grundlagenforschung und der Biotechnologie war *S. cerevisiae* der erste Eukaryot, dessen Genom sequenziert wurde.<sup>[172]</sup> Die ca. 6200 Gene inklusive der zugeordneten Funktion sind annotiert und katalogisiert in der *Saccharomyces Genome Database* ([www.yeastgenome.org](http://www.yeastgenome.org)) zu finden. Weiterhin wurden 95% des Genoms bereits systematisch deletiert und die generierten Stämme sind bei ATCC (*American Type Culture Collection*), Invitrogen oder EUROSCARF (*EUROpean Saccharomyces Cerevisiae Archive for Functional analysis*) kommerziell erhältlich. Die wachsende Anzahl an Datenbanken mit Informationen dieser Studien spiegelt das tiefe Verständnis seiner Genetik, Biochemie, Physiologie und Fermentationstechnik wieder.<sup>[173]</sup> Mehrere alternative Hefesysteme sind bekannt, die ebenfalls für die heterologe Expression von Genen Verwendung finden. Beispielhaft sind hier vier beschrieben. *Pichia pastoris* ist eine methylotrophe Hefe, mit deren Hilfe mittlerweile über 400 Proteine erfolgreich rekombinant produziert wurden.<sup>[174]</sup> Bei ihrer Fermentation erreicht sie extrem hohe Zelldichten ohne große Mengen von Ethanol oder Acetat zu akkumulieren, welche in hohen



Konzentrationen toxisch sind.<sup>[175]</sup> *Hansenula polymorpha* ist aufgrund der Fähigkeit, Xylose, Glukose und Cellobiose bei erhöhter Temperatur (48°C) zu vergären, für die Produktion von Bioethanol interessant.<sup>[176]</sup> Sowohl *Arxula* spp. als auch *Yarrowia lipolytica* (ehemals *Candida*) sind nicht nur thermotolerant bis zu einer Temperatur von 55°C, sondern zeichnen sich weiterhin durch einen temperaturabhängigen Dimorphismus aus. Während bei Inkubationstemperaturen bis 42°C eine Vermehrung durch Knospung initiiert wird, führen höhere Temperaturen zu Mycelbildung. Bedingt durch ein Morphismus-abhängiges Glykosylierungsmuster können sie daher für die heterologe Proteinproduktion sowohl O- als auch N-glykosylierter Proteine eingesetzt werden.<sup>[177]</sup> Für all diese Hefen wurden Vektoren zur genetischen Manipulation entwickelt.<sup>[176]</sup> Trotzdem fand im Gegensatz zu *S. cerevisiae* keines dieser Systeme Anwendung im *metabolic engineering* ganzer Stoffwechselwege. Der entscheidende Vorteil von *S. cerevisiae* gegenüber alternativen Hefen ist allerdings ihr Einsatz in vorherigen Untersuchungen zur Taxolbiosynthese. So wurden alle bisher bekannten Cytochrom-P450-Enzyme durch die Expression von *Taxus*-cDNS-Bibliotheken in *S. cerevisiae* entdeckt, bzw. für ihre Charakterisierung in *S. cerevisiae* expri-miert.<sup>[93][99][101][105][110][111][119][149]</sup> Die Bäckerhefe wurde daher auch im Rahmen dieser Arbeit als Expressionssystem ausgewählt.

*Saccharomyces cerevisiae* existiert in den zwei haploiden Paarungstypen (*Mating type*) MATa und MAT $\alpha$ . Beide sekretieren Pheromone (a-Faktor, bzw.  $\alpha$ -Faktor), die eine Zellfusion mit dem jeweils anderen Typen auslösen. Im Gegensatz zu den nicht modifizierten Pheromonen des Paarungstyps MAT $\alpha$ , enthält MATa zusätzliche strukturelle Komponenten, von denen eine als Farnesylgruppe identifiziert wurde.<sup>[178]</sup> Untersuchungen beider Paarungstypen auf heterologe Sesquiterpenproduktion zeigten eine doppelte Ausbeute beim Paarungstyp MATa gegenüber MAT $\alpha$ ,<sup>[126]</sup> sodass der MATa-Stamm CEN.PK2-1C<sup>[153]</sup> für die Durchführung der Dissertationsarbeiten ausgewählt wurde.

### **III.1.2 Selektion der Expressionsvektoren**

#### **III.1.2.1 Vektorsysteme**

Für die heterologe Expression von Genen in *Saccharomyces cerevisiae* stehen grundsätzlich drei Systeme zur Verfügung: die genomische Integration, die episomale Expression anhand eines zentromeren Vektors und die episomale Expression über ein 2  $\mu$ -basiertes Vektorsystem. Die Verwendung integrativer Vektoren verkörpert dabei die wohl stabilste Variante,<sup>[179]</sup> bei der die



fremde DNS über homologe Bereiche im linearisierten Vektor in das Hefegenom eingefügt wird<sup>[180]</sup>. Mittlerweile wurden bereits zahlreiche Vektorsysteme für die genomische Integration entwickelt, die in ihrer Selektionsmöglichkeit, dem Integrationslokus, ihrem Promotor und Zahl der Integrationsereignisse variieren.<sup>[181-186]</sup> Gleichzeitig bedeutet die Integration aber auch einen hohen Grad an Inflexibilität, was den Austausch einzelner Gene betrifft, sodass sie im Rahmen dieser Arbeit keine Verwendung fand. Die Expression anhand episomaler Systeme kann auf (*autonomous replication site*) ARS-Vektoren basieren, die innerhalb der Zelle in multipler Kopienzahl vorhanden sind (1-20). Durch eine ineffiziente Verteilung auf die Tochterzellen sind sie allerdings sehr instabil. Ohne Selektion gehen daher rund 20 % von ihnen innerhalb einer Generation verloren.<sup>[187]</sup> Durch Addition einer zentromeren Sequenz (CEN) können ARS-Vektoren zwar stabilisiert werden, dies verringert ihre Kopienzahl allerdings auch auf 1-2.<sup>[188]</sup> Die wohl meist verwendeten episomalen Expressionsvektoren basieren auf der hefeeigenen 2  $\mu$ -Sequenz. 2  $\mu$  ist ein 6,3 kb großes Plasmid, das in den meisten *S. cerevisiae*-Stämmen zu finden ist und in 60-100 facher Kopienzahl vorliegt. Seine Stabilität wird durch den Replikationsursprung, der sich wie eine typische ARS verhält, und dem *STB*-Lokus gewährleistet. Trotz der weiten Verbreitung ist keine Funktion bekannt. Die stabile Vererbung basiert auf seiner Fähigkeit, sich nach ungleicher Verteilung auf die Tochterzelle bis zur richtigen Kopienzahl zu amplifizieren.<sup>[189]</sup> Der einfachste 2  $\mu$ -basierte Vektor besitzt daher die *ORI-STB*-Sequenz, einen Selektionsmarker und bakterielle Plasmidsequenzen zur dortigen Replikation und Selektion. Je nach Selektionsmarker erreicht der 2  $\mu$ -Vektor 50-400 Kopien innerhalb der Hefezelle<sup>[179]</sup> und wurde daher als grundlegendes System für den Einsatz in dieser Arbeit ausgewählt.

### III.1.2.2 Promotorwahl

Neben einer hohen Kopienzahl war weiterhin entscheidend, dass der selektierte Vektor bereits die entsprechenden Promotor-, bzw. Terminationssequenzen aufweist und eine unkomplizierte Klonierung ermöglicht. Die beiden Eigenschaften werden von den pVV-Vektoren<sup>[154]</sup> verbunden, die 2  $\mu$ -basierte Vektorsysteme mit einer hohen Kopienzahl und Gateway-Kompatibilität darstellen. Sie sind von EUROSCARF sowohl mit konstitutiven Promotoren aus der Glykolyse als auch mit induzierbaren Promotoren erhältlich. Alle bisherigen heterologen Expressionen von Taxolbiosynthesegenen in *S. cerevisiae* erfolgten mithilfe von Galaktose-induzierbaren Promotoren.<sup>[99][105][119][149]</sup> Auch für die rekombinante Produktion anderer terpenoider Substanzen wie Sesquiterpenen<sup>[126]</sup> und Pregnenolon, bzw. Progesteron<sup>[170]</sup> fanden sie Verwendung. Galaktose-induzierbare Promotoren vertreten die wohl stärksten regulierbaren Promotoren in der Bäckerhefe

und führen bei Induktion zu entsprechender mRNA, die 1% der Gesamt-RNS ausmacht.<sup>[190]</sup> Sie finden vor allen Dingen Verwendung, wenn die Genprodukte durch ihre katalytische Aktivität zur Akkumulation toxischer Substanzen führen oder die gewünschte Verbindung ausschließlich in bestimmten Wachstumsphasen synthetisiert werden soll. Allerdings ist ihre Induktion aufgrund der Repression durch Glukose stets mit der Eliminierung von Restglukose verbunden. Für einen genauen Induktionszeitpunkt bedeutet dies einen Medienwechsel, was mit wachsendem Maßstab weniger praktikabel wird. Neben dem zusätzlichen Zeitaufwand bedeutet der Einsatz von Galaktose eine kostspielige C-Quelle. Dies wird durch die Verwendung von konstitutiven Glykolysepromotoren umgangen. Mit 5% der Gesamt-RNS führen sie weiterhin zu der wohl stärksten Expression in *S. cerevisiae*.<sup>[179]</sup> Die Produkte der im Rahmen dieser Arbeit exprimierten Gene führen zu keiner Akkumulation toxischer Substanzen und werden für die sequenzielle Umsetzung von Substraten in großen Mengen benötigt. Die Verwendung der 2  $\mu$ -Vektoren pVV214 und pVV200 mit dem konstitutiven Phosphoglyceratkinase-Promotor stellt daher eine kostengünstige Alternative zu entsprechenden Vektoren mit induzierbaren Promotoren dar und verkörpert zusätzlich eine alternative Strategie des *metabolic engineering* von Hefen zur rekombinanten Produktion von Taxanen.

### III.2 Generierung eines Taxadien produzierenden Hefestamms

Das Enzym Taxadiensynthase (TDS) katalysiert den ersten Schritt der Taxolbiosynthese, bei dem Geranylgeranyldiphosphat zu Taxadien zyklisiert wird.<sup>[82]</sup> Als plastidäres Protein enthält es eine für den Transport in die Plastiden benötigte N-terminale Signalsequenz. Bedingt durch den nur eingeschränkt bestimmbaren Übergangsbereich zum prozessierten Portein, wurden sowohl das Volllängeprotein als auch verschiedene verkürzte Varianten auf ihre Aktivität untersucht (Abbildung I.14). Die um 60 Aminosäuren verkürzte TDS stellte sich nicht nur als die löslichste und stabilste Variante heraus, sondern zeigte auch die beste, mit dem pflanzlichen Enzym vergleichbare, katalytische Aktivität.<sup>[91]</sup> Für die heterologe Expression in *S. cerevisiae* wurde daher das für die TDS dieser Länge kodierende Gen in den pVV214 inseriert. Als Matrize für die Amplifikation des Taxadiensynthasegens diente dabei der entsprechende cDNA-Klon aus *Taxus chinensis*. Die verwendeten Primer 1 und 2 (Tabelle II.3) ermöglichten durch Anhängen entsprechender Rekombinationsstelle an die Gensequenz die Klonierung mittels Gateway-Technologie (II.2.6.2). Mit dem so generierten Vektor pVV214::*TDS-Tc* (Tabelle III.1) wurde der *S. cerevisiae* Stamm CEN.PK2-1C transformiert und der rekombinante Hefestamm CEN 1

generiert. Alle im Rahmen der Generierung eines Taxadien produzierenden Stamms konstruierten Hybridplasmide sind in Tabelle III.1 aufgelistet.

**Tabelle III.1: Generierte Hybridplasmide**

Plasmid	Relevanter Genotyp	Vektoranteil
pVV214:: <i>TDS-Tc</i>	Taxadiensynthasegen aus <i>Taxus chinensis</i> ohne Singalpepti	pVV214
pVV200:: <i>GGPPS-Tc</i>	GGPP-Synthasegen aus <i>Taxus chinensis</i>	pVV200
pVV200:: <i>UPC2.1-Sc</i>	Gen für punktmutierten Transkriptionsfaktor UPC2 aus <i>S. cerevisiae</i>	pVV200
pRS313:: <i>UPC2.1-Sc</i>	Gen für punktmutierten Transkriptionsfaktor UPC2 aus <i>S. cerevisiae</i>	pRS313
pVV200:: <i>tHMGRI-Sc</i>	verkürztes HMG-CoA-Reduktasegen 1 aus <i>S. cerevisiae</i>	pVV200
pRS315:: <i>tHMGRI-Sc</i>	verkürztes HMG-CoA-Reduktasegen 1 aus <i>S. cerevisiae</i>	pRS315
pVV200:: <i>GGPPS-Sa</i>	GGPP-Synthasegen aus <i>Sulfolobus acidocaldarius</i>	pVV200
pVV214:: <i>TDSco</i>	TDS-Gen, codon-optimiert für die Expression in <i>S. cerevisiae</i>	pVV214

**TDS:** Taxadiensynthase; **GGPP:** Geranylgeranyldiphosphat; **HMG-CoA:** 3-hydroxy-3-methylglutaryl-Coenzym A

Die für eine rekombinante Taxadiensynthese benötigten Isoprenoidvorstufen werden in Hefen über den Mevalonatweg synthetisiert, von *Saccharomyces cerevisiae* allerdings hauptsächlich für die Biosynthese von Ergosterol eingesetzt. Diese Verbindung sorgt neben anderen Funktionen vorwiegend für die Stabilität und Fluidität der Zellmembran.<sup>[191-193]</sup> Biogenetisch leitet sich Ergosterol von Farnesyldiphosphat FPP ab, welches mit einem zweiten FPP-Molekül zu Squalen, der ersten spezifischen Vorstufe des Ergosterols, kondensiert. Es ist daher nicht erstaunlich, dass FPP-abgeleitete Substanzen rekombinant bereits erfolgreich in *Saccharomyces cerevisiae* produziert wurden, ohne den hefeeigenen Primärmetabolismus zu verändern.<sup>[170][171]</sup> Wie bereits von DeJong beschrieben, ist der endogene Pool an Geranylgeranyldiphosphat (GGPP) in *S. cerevisiae* dagegen sehr gering.<sup>[149]</sup> Zur effizienten Bereitstellung des für die Synthese von Taxadien benötigten GGPPs wurde daher das Gen für die GGPP-Synthase (GGPPS) aus *Taxus chinensis* coexprimiert. Entsprechend der Klonierung des *TDS* wurde für die heterologe Expression in Hefe auch *GGPPS* mittels Gatewaytechnologie (II.2.6.2) aus dem entsprechenden *Taxus* cDNA-Klon in den Destinationsvektor inseriert (Primer 2 und 3, Tabelle II.3). Mit dem entstandenen Vektor pVV200::*GGPPS-Tc* (Tabelle III.1) wurde der *TDS*-exprimierende Hefestamm CEN 1 transformiert und damit der Stamm CEN 2 generiert. CEN 1 und CEN 2, sowie alle weiteren in III.2 beschriebenen Hefestämme sind mit den jeweils transformierten Vektoren in Tabelle III.2 aufgelistet.

**Tabelle III.2: Generierte Hefestämme**

Stamm	Transformierte Vektoren
CEN 1	pVV214:: <i>TDS-Tc</i>
CEN 2	pVV214:: <i>TDS-Tc</i> ; pVV200:: <i>GGPPS-Tc</i>
CEN 3	pVV214:: <i>TDS-Tc</i> ; pVV200:: <i>GGPPS-Tc</i> ; pRS315:: <i>tHMGR1-Sc</i>
CEN 4	pVV214:: <i>TDS-Tc</i> ; pVV200:: <i>GGPPS-Tc</i> ; pRS315:: <i>tHMGR1-Sc</i> ; pRS313:: <i>UPC2.1-Sc</i>
CEN 5	pVV214:: <i>TDS-Tc</i> ; pVV200:: <i>GGPPS-Sa</i> ; pRS315:: <i>tHMGR1-Sc</i> ; pRS313:: <i>UPC2.1-Sc</i>
CEN 6	pVV214:: <i>TDSco</i> ; pVV200:: <i>GGPPS-Tc</i> ; pRS315:: <i>tHMGR1-Sc</i> ; pRS313:: <i>UPC2.1-Sc</i>
CEN 7	pVV214:: <i>TDSco</i> ; pVV200:: <i>GGPPS-Sa</i> ; pRS315:: <i>tHMGR1-Sc</i> ; pRS313:: <i>UPC2.1-Sc</i>

***TDS-Tc***: Taxadiensynthasegen aus *Taxus chinensis*; ***GGPPS-Tc***: Geranylgeranyldiphosphatsynthasegen aus *T. chinensis*; ***tHMGR1-Sc***: Gen für HMG CoA-Reduktase Isozym 1 aus *Saccharomyces cerevisiae*; ***UPC2.1-Sc***: punktmutierte Version des Transkriptionsfaktors UPC2 aus *S. cerevisiae*; ***GGPPS-Sa***: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; ***TDSco***: für die Expression in *S. cerevisiae* codon-optimiertes Taxadiensynthasegen

Wie in den prinzipiellen Überlegungen bereits ausgeführt, basiert die erfolgreiche heterologe Produktion von Taxadien in *S. cerevisiae* auf einer ausreichenden Menge an Vorstufen aus dem Hefe-Primärstoffwechsel. Um diese zu gewährleisten, wurde der hefeeigene Mevalonatweg modifiziert. Die Endprodukte dieses Synthesewegs sind IPP und DMAPP, aus denen die jeweiligen Grundgerüste aller Isoprenoide synthetisiert werden. Die Regulation des Mevalonatwegs erfolgt dabei in der Umsetzung von HMG-CoA zu Mevalonat durch die HMG-CoA-Reduktase (Abbildung I.8). Um auch im Falle des Taxadien produzierenden Stammes eine Regulierung der Reduktase und damit der Produktion der Isoprenoidvorstufen zu umgehen, wurde die verkürzte Version des Gens (*tHMGR1-Sc*) exprimiert, die ausschließlich für die katalytische Domäne des Reduktase-Isozyms 1 kodiert. Die Expression dieser *N*-terminal um die membranassoziierte regulatorische Domäne verkürzte Version hatte bereits zu einer 20-fachen Steigerung der intrazellulären Reduktaseaktivität geführt.<sup>[69]</sup>

Der verwendete Hefestamm CEN.PK2-1C weist vier Auxotrophien auf, die zur Selektion transformierter Vektoren benutzt werden können: Uracil, Tryptophan, Histidin und Leucin (Tabelle II.1). Durch die Transformation mit pVV214::*TDS-Tc* und pVV200::*GGPPS-Tc* komplementiert der generierte Stamm CEN 2 bereits Uracil (pVV214) und Tryptophan (pVV200). Zum Zeitpunkt der beschriebenen Klonierungen waren keine Gateway-kompatiblen Vektoren mit den verbleibenden Selektionsmarkern Histidin und Leucin verfügbar. Den kommerziell erhältlichen Vektoren mit den benötigten Selektionsmöglichkeiten fehlten andererseits die Regulationseinheiten in Form von Promotor und Terminator. Um die Gateway-Kompatibilität und den konstitutiven Glykolyse-Promotor der pVV-Vektoren mit den Selektionsmarkern der pRS-Vektoren (Tabelle II.2)

zu kombinieren, wurden die Klonierungen aller weiteren Gene in zwei Schritten durchgeführt. Nach Gateway-Rekombination (II.2.6.2) in einen der pVV-Vektoren wurden die Gene inklusive des Promotors und Terminators aus dem Vektor amplifiziert und dabei mit Schnittstellen versehen, die eine Klonierung über Restriktion/Ligation (II.2.5) des gesamten Fragments in einen pRS-Vektor ohne eigene Regulationseinheiten, aber mit dem benötigten Selektionsmarker ermöglichen (Abbildung III.1).



**Abbildung III.1: Klonierungsstrategie weiterer Gene am Beispiel von *tHMGR1*.** Für die Expression weiterer Gene neben *TDS* und *GGPPS* wurden diese erst in den Gateway-kompatiblen pVV200-Vektor mit Promotor- und Terminator inseriert. Um die Limitierung der Selektionsmarker dieser Vektoren zu umgehen, wurden die Gene anschließend inklusive Regulationseinheiten aus pVV200 amplifiziert und über die so angehängten Restriktionsschnittstellen *SacII* und *NotI* in den ebenso verdautes pRS-Vektoren mit variierenden Selektionsmarkern ligiert.

Die Klonierung erfolgte durch Amplifikation des hefeeigenen *tHMGR1*-Gens mit Primern 5 und 6 (Tabelle II.3) und anschließender Gateway-Rekombination in pVV200. Für die Expression von *tHMGR1-Sc* in pRS315 wurde das Fragment *PGK*-Promotor-*tHMGR1-Sc*-*CYC1*-Terminator mit Primern 7 und 8 (Tabelle II.3) aus dem generierten pVV200::*tHMGR1-Sc* amplifiziert und anhand

der angehängten Schnittstellen *NotI* und *SacII* nach Restriktion mit den entsprechenden Endonukleasen in den ebenso geschnittenen Vektor pRS315 ligiert. Mit dem generierten Hybridplasmid pRS315::*tHMGR1-Sc* (Tabelle III.1) wurde CEN 2 transformiert und damit der Hefestamm CEN 3 generiert (Tabelle III.2).

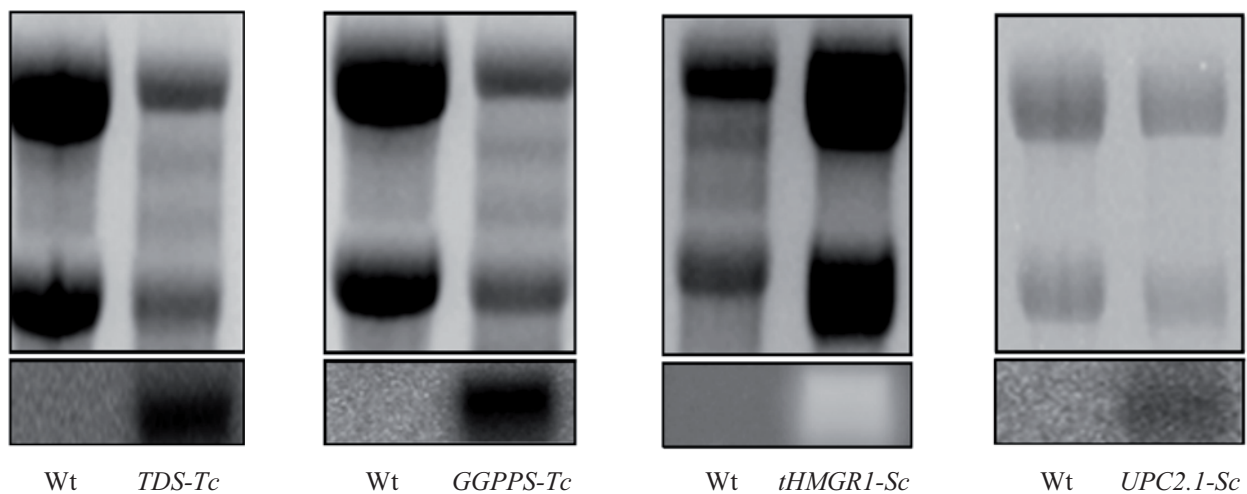
Die rekombinante Produktion des Diterpens Taxadien steht in Konkurrenz zur hefeeigenen Sterolproduktion, da beide Substanzen bis hin zum Farnesyldiphosphat (FPP) den gleichen Biosyntheseweg aufweisen (Abbildung I.1). Die Synthese von Ergosterol ist für die Hefe ein physiologisch aufwendiger Prozess, der aus 25 Reaktionsschritten besteht und je nach Modifikation mindestens 10 ATP und 16 NADPH-Äquivalente verbraucht.<sup>[191]</sup> Die Aufnahme externen Sterols könnte diese nicht nur einsparen, sondern auch einen verstärkten Kohlenstofffluss von FPP in Richtung GGPP begünstigen. Im Gegensatz zu anderen Mikroorganismen, die die biosynthetische Kontrolle von Nährstoffen mit deren exogenen Verfügbarkeit abstimmen,<sup>[194]</sup> synthetisiert *Saccharomyces cerevisiae* unter aeroben Bedingungen große Mengen an Ergosterol, ohne auf externe Quellen zuzugreifen.<sup>[195]</sup> Dabei ist Sauerstoff nicht nur für die Synthese von Ergosterol selbst, sondern auch für diejenige von Fettsäuren nötig, mit denen Sterole in Hefe verestert werden<sup>[196]</sup>. Durch die Mutation eines unter anaeroben Bedingungen normalerweise nicht zur Sterolaufnahme fähigen Wildtypstamms<sup>[195][197]</sup> fanden Lewis und Mitarbeiter anhand zugefütterten <sup>14</sup>C-markierten Sterols einen Stamm, der in seiner Sterolaufnahme (*uptake control* UPC) verändert war.<sup>[194]</sup> Genauere Untersuchungen ergaben, dass *upc2* für einen Transkriptionsfaktor kodiert, für den ein Drittel aller bei Hypoxie induzierten Gene mindestens eine potenzielle Bindungsstelle aufweisen.<sup>[198]</sup> Es ist bekannt, dass Hefen bei Sauerstoffmangel die Zusammensetzung ihrer Zellwand verändern, welche unter anderem für die selektive Aufnahme von Makromolekülen verantwortlich ist.<sup>[199]</sup> Es erstaunt daher nicht, dass die Expression der punktmutierten Version *upc2.1* zur Fähigkeit der Sterolaufnahme auch unter aeroben Bedingungen führte.<sup>[200]</sup> UPC2-1 unterscheidet sich dabei von seinem Wildtyppendant ausschließlich durch einen einzelnen Aminosäureaustausch in der C-terminalen regulatorischen Domäne,<sup>[201]</sup> welcher zu einer effizienteren Aktivierung der entsprechenden Gene führt.<sup>[202]</sup>

Für einen durch die Aufnahme externen Sterols ermöglichten besseren Kohlenstofffluss in Richtung des Diterpens Taxadien sollte daher auch *upc2.1* in CEN 3 exprimiert werden. Die Punktmutation wurde mittels SOE-PCR (II.2.2.2) in das aus genomischer DNS amplifizierte Gen eingeführt und das Produkt über angehängte Gatewaysequenzen in pVV200 rekombiniert (Primer 9-12, Tabelle II.3). Wie schon *tHMGR-Sc* wurde auch *UPC2.1-Sc* mitsamt Promotor- und Terminationssequenz aus pVV200::*UPC2.1-Sc* amplifiziert (Primer 13 und 14) und via



Restriktion / Ligation in den Histidin-komplementierenden Vektor pRS313 kloniert. Durch Transformation mit dem generierten Hybridplasmid pRS313::*UPC2.1-Sc* (Tabelle III.1) in CEN 3 (Tabelle III.2) entstand so der Hefestamm CEN 4 (Tabelle III.2).

Alle klonierten Gene wurden anhand eines Northernblots auf ihre erfolgreiche Expression getestet. Abbildung III.2 gibt den autoradiographischen Nachweis wieder, visualisiert auf einer Phosphorplatte (II.3.4.4). Die Sequenzen der zur Detektion eingesetzten Oligonukleotide sind in Tabelle II.5 aufgelistet.



**Abbildung III.2:** Northern blot-Analyse der in CEN1-4 exprimierten Gene. **Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS. **Unten:** Autoradiogramm der spezifisch gebundenen, radioaktiv markierten DNS-Sonde. **Wt:** CEN.PK2-1C Wildtyp; ***TDS-Tc*:** Taxadiensynthase aus *Taxus chinensis*; ***GGPPS-Tc*:** Geranylgeranyldiphosphatsynthase aus *T. chinensis*; ***tHMGR1-Sc*:** HMG-CoA-Reduktase Isozym 1 aus *Saccharomyces cerevisiae*; ***UPC2.1-Sc*:** Punktmutierte Version des Transkriptionsfaktors UPC2 aus *Saccharomyces cerevisiae*

Trotz Nachweis einer erfolgreichen Transkription mittels Northern blot konnte via Western blot-Analyse eine Translation der *TDS-Tc*- bzw. *GGPPS-Tc*-mRNA nicht bestätigt werden. Der indirekte Nachweis wurde hier durch Detektion der jeweiligen Katalyseprodukte anhand von GC/MS-Messungen (II.3.5.1) erbracht. Die so ermittelten Produktmengen sind in Tabelle III.3 zusammengefasst. Bei rekombinanter Expression von *TDS-Tc* konnte Taxadien zwar detektiert, aber bedingt durch die geringe Menge nicht quantifiziert werden. Wurde *GGPPS-Tc* ebenfalls exprimiert, war bei einer mit dem Wildtyp vergleichbaren Menge an gebildetem GGOH ( $235 \pm 12 \mu\text{g/l}$ ) eine Taxadienproduktion von  $204 \pm 1 \mu\text{g/l}$  nachweisbar. Diese wurde durch die zusätzliche Expression von *tHMGR-Sc* um 50% auf  $306 \pm 3 \mu\text{g/l}$  an Taxadien gesteigert, während die GGOH-Produktion nur einen leicht erhöhten Wert von  $253 \pm 34 \mu\text{g/l}$  ergab. Bei Ergänzung der

rekombinanten Gene durch *UPC2.1-Sc* erhöhte sich die GGOH-Quantität erneut geringfügig auf  $283 \pm 32 \mu\text{g/l}$ , während die gemessene Taxadienkonzentration mit  $306 \pm 1 \mu\text{g/l}$  gleich blieb.

**Tabelle III.3: Untersuchung generierter Hefestämme auf Produktquantität mittels GC/MS**

Stamm	Überexprimierte Gene	Taxadien [ $\mu\text{g/l}$ ]	GGOH [ $\mu\text{g/l}$ ]
CEN wt	keine	nicht detektierbar	$231 \pm 3$
CEN 1	<i>TDS-Tc</i>	nicht quantifizierbar	$229 \pm 8$
CEN 2	<i>TDS-Tc; GGPPS-Tc</i>	$204 \pm 1$	$235 \pm 12$
CEN 3	<i>TDS-Tc; GGPPS-Tc; tHMGR1-Sc</i>	$306 \pm 3$	$253 \pm 34$
CEN 4	<i>TDS-Tc; GGPPS-Tc; tHMGR1-Sc; UPC2.1-Sc</i>	$306 \pm 1$	$283 \pm 32$

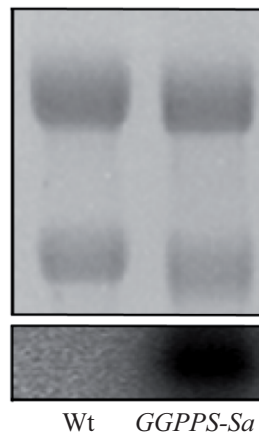
**Wt:** CEN.PK2-1C Wildtyp; ***TDS-Tc*:** Taxadiensynthase aus *Taxus chinensis*; ***GGPPS-Tc*:** Geranylgeranyldiphosphat-Synthase aus *Taxus chinensis*; ***tHMGR1-Sc*:** HMG-CoA-Reduktase Isozym 1 aus *Saccharomyces cerevisiae*; ***UPC2.1-Sc*:** Punktmutierte Version des Transkriptionsfaktors UPC2 aus *Saccharomyces cerevisiae*

Um die GGPP-Produktivität und dadurch bedingt auch die Taxadien-Ausbeute weiter zu steigern, wurde die pflanzliche GGPPS mit dem Analogon aus dem Archaeobakterium *Sulfolobus acidocaldarius* (GGPPS-Sa) ausgetauscht. Archaeobakterien stellen neben Eubakterien und Eukaryoten die dritte Lebensdomäne dar und sind bekannt für ihr Vorkommen unter extremen Umweltbedingungen. Je nach Habitat werden sie in Halophile (hohe Salzkonzentration), Thermophile (hohe Temperaturen), Thermoacidophile (hohe Temperaturen, niedriger pH) und Methanogene (strikt anaerob) eingeteilt.<sup>[203]</sup> Die Besetzung dieser extremen Lebensräume wird ihnen durch die Fähigkeit ermöglicht, die Permeabilität ihrer Membran den äußeren Bedingungen entsprechend zu verändern und so den osmotischen Druck innerhalb der Zelle, sowie die Nährstoffaufnahme und eine Energiegewinnung durch transmembrane elektrochemische Gradienten zu gewährleisten.<sup>[204]</sup> Im Gegensatz zu allen anderen bekannten lebenden Organismen, deren Membran auf der Veresterung von Alkoholen mit Fettsäuren basiert, bestehen archaebakterielle Membranen aus Etherverbindungen, deren verzweigte Kohlenwasserstoffketten nicht von Fettsäuren, sondern von den isoprenoiden Farnesyl- bzw. Geranylgeranylderivaten abstammen.<sup>[205]</sup> Obwohl auch Eubakterien und Eukaryoten ihre Membranen durch verschiedene Sättigungsgrade der Fettsäuren modifizieren können, sind sie aufgrund deren Grundstruktur und der daraus resultierenden Permeabilitätsgrenze nur über einen kleinen Temperatur- und pH-Bereich fähig, einen Protonengradienten aufrecht zu erhalten. Die archaebakteriellen Etherverbindungen hingegen besitzen eine größere chemische Stabilität. Verzweigungen innerhalb der Isoprenmoleküle ermöglichen zudem eine dichtere Packung, was zu einer geringeren Ionenpermeabilität führt.<sup>[206]</sup> Die chemische Zusammensetzung archaebakterieller Membranlipide ist so einzigartig und vielseitig, dass sie neben der rRNS-Sequenzierung nicht nur als taxonomischer Marker dient,



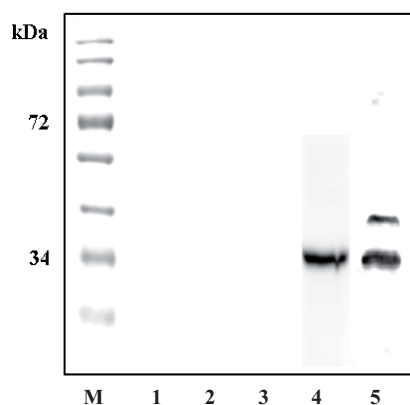
sondern ihre Fluiditäts- und Permeabilitätseigenschaften auch als Vorbild für künstliche Membranen mit Langzeit- und Thermostabilität dienen. So wurde mit sogenannten Archaeosomen nicht nur eine neue Generation von Liposomen für den Einsatz in der Pharmakotherapie entwickelt, sondern auch diverse Anwendungen im Bereich der Impfstoffentwicklung patentiert.<sup>[207]</sup>

Während Diterpene in *Saccharomyces cerevisiae* hauptsächlich für die Dolichol- und Ubichinonsynthese sowie zur Prenylierung von Proteinen eingesetzt werden, spielen sie bei Archaeobakterien also zusätzlich eine Schlüsselrolle bei der Membranbiosynthese. Auch in Archaeobakterien wird GGPP durch die Kondensation von DMAPP und IPP über die Zwischenprodukte Geranyldiphosphat (GPP) und FPP synthetisiert. Im Gegensatz zu ihren eubakteriellen, bzw. eukaryotischen Gegenstücken, bei denen jeder einzelne Reaktionsschritt von einer spezifischen Prenyltransferase katalysiert wird,<sup>[208]</sup> wurde in Archaeobakterien ein bifunktionales Enzym gefunden, welches den gesamten Syntheseweg von IPP und DMAPP zu GGPP katalysiert.<sup>[209][210]</sup> Die Kondensation der Vorstufen IPP und DMAPP zu GGPP durch die katalytische Aktivität eines einzigen Enzyms könnte eine effektivere Umsetzung der entstandenen Intermediate zum Endprodukt ermöglichen, sodass durch die Aktivität einer archaeobakteriellen bifunktionalen GGPPS die nötigen Substrate für eine effektivere rekombinante Taxadienproduktion bereitgestellt werden sollten. Die bifunktionale GGPP-Synthase aus *Sulfolobus acidocaldarius*<sup>[210a]</sup> wurde bereits erfolgreich in einem eukaryotischen System exprimiert,<sup>[211]</sup> sodass das entsprechende Gen *GGPPS-Sa* für die heterologe Expression in *Saccharomyces cerevisiae* ausgewählt wurde. Die Klonierung von *GGPPS-Sa* erfolgte analog zu der bereits für *GGPPS-Tc* beschriebenen. Entsprechend seiner außergewöhnlichen Kultivierungsbedingungen (75-80°C, pH2-3, strikt anaerob<sup>[212]</sup>) wurde *Sulfolobus acidocaldarius* in DSM-Medium 88 (II.1.2.5) bei 75°C inkubiert und daraus isolierte genomische DNS in die Amplifikationsreaktion eingesetzt (Primer 13 und 14, Tabelle II.3). Mit dem generierten Hybridplasmid pVV200::*GGPPS-Sa* (Tabelle III.1) wurde anschließend der *TDS* exprimierende Stamm CEN 1 transformiert. Durch das zusätzliche Einbringen der Plasmide pRS315::*tHMGR1-Sc* und pRS313::*UPC2.1-Sc* entstand so der Hefestamm CEN 5 (Tabelle III.2). Die problemlose Transkription des archaeobakteriellen Gens in *Saccharomyces cerevisiae* wurde mittels Northernblot-Analyse bestätigt (Abbildung III.3).



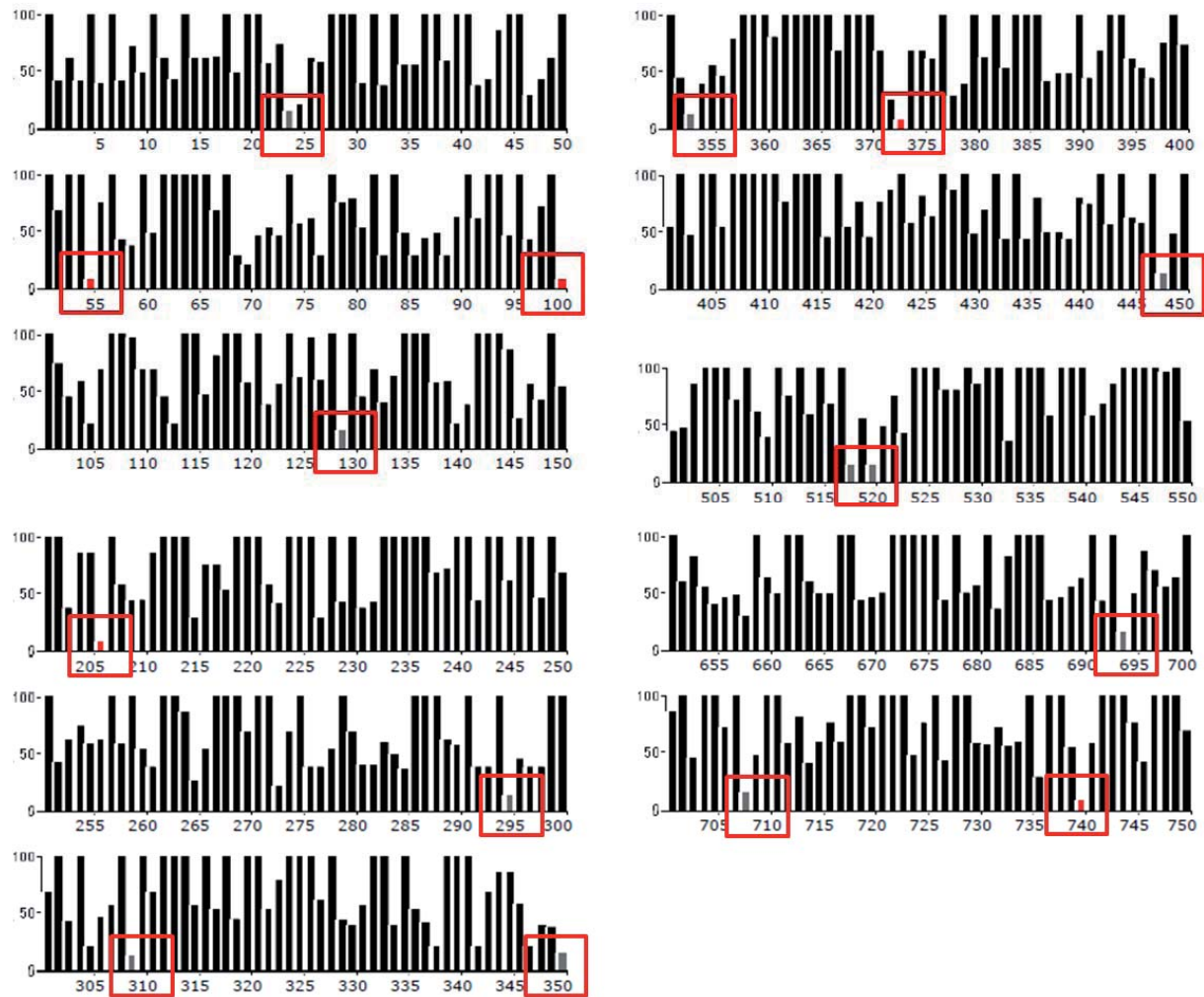
**Abbildung III.3:** Northernblot der in *Saccharomyces cerevisiae* exprimierten GGPP-Synthase aus *Sulfolobus acidocaldarius* (GGPPS-*Sa*). **Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10 µg. **Unten:** Autoradiogramm der spezifisch gebundenen, radioaktiv markierten DNS-Sonde. **Wt:** CENPK2-1C Wildtyp

Der zur Überprüfung einer erfolgreichen Translation synthetisierter mRNA durchgeführte Westernblot ist in (Abbildung III.4) dargestellt. Das Genprodukt der hefeeigenen, homolog exprimierten *tHMGR1-Sc* konnte anhand einer Proteinbande richtiger Größe (38 kDa) identifiziert werden. Die Produkte der pflanzlichen Gene *TDS-Tc* und *GGPPS-Tc* waren trotz ihrer durch Sequenzierung bestätigten, korrekten Basenabfolge (inklusive Histidinmarker) und nachgewiesener Transkription (Abbildung III.2) hingegen nicht detektierbar. Eine Translation konnte hier ausschließlich indirekt über die Detektion gebildeter Katalyseprodukte mittels GC/MS nachgewiesen werden (Tabelle III.3). Wurde die pflanzliche *GGPPS-Tc* allerdings durch das Archaeengen *GGPPS-Sa* ersetzt, zeigte die Analyse der entsprechenden Proteinfraction ein positives Signal in Form einer Bande korrekter Größe (34 kDa).



**Abbildung III.4:** Westernblot von 40 µg Gesamtprotein generierter Taxadien Produktionsstämme. **M:** Größenstandard (Fermentas); **1:** CEN.PK2-1C Wildtyp; **2:** CEN 1 (pVV214::*TDS-Tc*); **3:** CEN 2 (pVV214::*TDS-Tc*; pVV200::*GGPPS-Tc*); **4:** CEN (pVV214::*TDS-Tc*; pVV214::*GGPPS-Sa*); **5:** CEN 5 ((pVV214::*TDS-Tc*; pVV214::*GGPPS-Sa*; pRS315::*tHMGR1-Sc*)

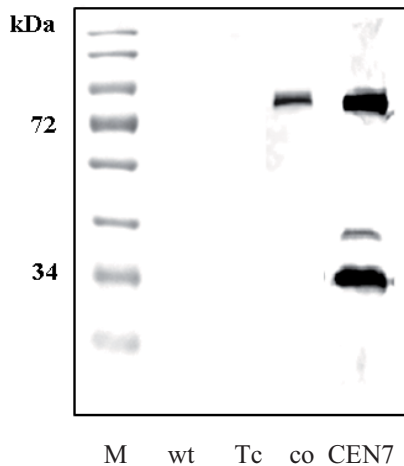
Ein Abgleich der pflanzlichen *TDS-Tc* mit der Codonverteilung von *Saccharomyces cerevisiae* ergab das Vorkommen mehrerer Arginincodons innerhalb der Proteinsequenz, die in der Hefe selten vorkommen und für die daher nur geringe Mengen an entsprechender tRNA zur Verfügung stehen (Abbildung III.5A).



**Abbildung III.5: Relative Anpassung der Codonverteilung der Taxadiensynthese aus *Taxus chinensis* an die von *Saccharomyces cerevisiae*.** **Rote Kästchen:** Markierter Bereich, in dem der Schwellenwert unterschritten wird; **Grau:** Wert unter 20 %; **Rot:** Wert unter 10 % (www.gcua.de)

Um die Translation der Taxadiensynthese zu verbessern, wurde daher ein für *Saccharomyces cerevisiae* codon-optimiertes Gen verwendet (Abbildung III.5). Das von der Firma Geneart synthetisierte Gen wurde mittels Gateway-Technologie in pVV214 inseriert. Der generierte Vektor pVV214::*TDSco* (Tabelle III.1) wurde in den Hefestamm CEN.PK2-1C (Tabelle II.1) transformiert und durch sequenzielles Einbringen der Vektoren pVV200::*GGPP-Tc*, pRS315::*tHMGR1-Sc* und pRS313::*UPC2.1-Sc* der Stamm CEN 6 (Tabelle III.2) generiert. Um die Auswirkungen von *GGPPS-Sa* auf die Taxadienproduktion genauer analysieren zu können, wurde des Weiteren der Vektor pVV200::*GGPPS-Tc* gegen pVV200::*GGPPS-Sa* ausgetauscht und der Stamm CEN 7 generiert (Tabelle III.2).

Abbildung III.6 zeigt die anhand eines Westernblots überprüfte Translation bei Expression des pflanzlichen Taxadiensynthasegens im Vergleich mit jener des synthetischen Gens.



**Abbildung III.6: Westernblot von 40 µg Gesamtprotein generierter Taxadien Produktions-stämme.**  
**M:** Größenstandard (NEB); **wt:** CEN.PK2-1C Wildtyp; **Tc:** CEN 1 (pVV214::TDS-Tc); **co:** CEN (pVV214::TDSco); **CEN 7:** CEN (pVV214::TDSco; pVV200::GGPPS-Sa; pRS315::tHMGR1-Sc; pRS313::UPC2.1-Sc)

Die durchgeführten Westernblot-Analysen von 40 µg Gesamtprotein der neuen Stämme zeigten eine deutlich verbesserte Translation der synthetischen Taxadiensynthase, die mit einer Proteinbande der richtigen Größe (90 kDa) sichtbar wurde. Die generierten Hefestämme mit synthetischer TDS wurden mittels GC/MS (II.3.5.1) auf ihre Produktion an Taxadien und GGOH überprüft. Tabelle III.4 gibt die gemessenen Werte im Vergleich zu den das pflanzliche GGPPS-Gen exprimierenden Stämmen wieder.

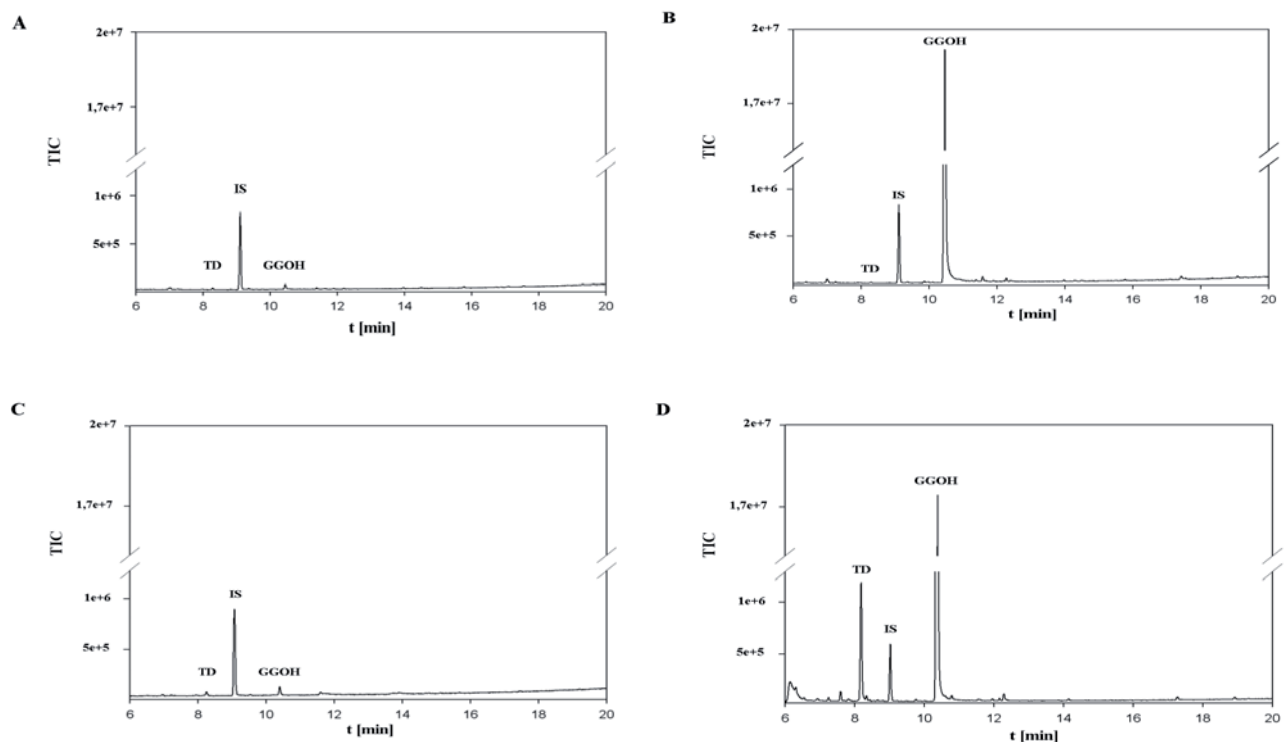
**Tabelle III.4: Untersuchung optimierter Hefestämme auf Produktquantität mittels GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]
CEN wt	keine	n.d.	0,23±0,003
CEN 4	<i>TDS-Tc</i> ; <i>GGPPS-Tc</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	0,31±0,001	0,28±0,03
CEN 5	<i>TDS-Tc</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	0,32±0,003	27,6±3,4
CEN 6	<i>TDSco</i> ; <i>GGPPS-Tc</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	0,22±0,02	0,29±0,01
CEN 7	<i>TDSco</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	8,7 ± 0,85	33,1±5,6

**Wt:** CEN.PK2-1C Wildtyp; **TDSco:** für die Expression in *S. cerevisiae* codon-optimierte Taxadiensynthase; **GGPPS-Tc:** Geranylgeranyldiphosphatsynthase aus *Taxus chinensis*; **GGPPS-Sa:** Geranylgeranyldiphosphatsynthase aus *Sulfolobus acidocaldarius*; **tHMGR1-Sc:** HMG-CoA-Reduktase Isozym 1 aus *S. cerevisiae*; **UPC2.1-Sc:** Punktmutierte Version des Transkriptionsfaktors UPC2 aus *Saccharomyces cerevisiae*; **n.d.:** nicht detektierbar

Ein Austausch des pflanzlichen Taxadiensynthasegens gegen eine für die Expression in *S. cerevisiae* codon-optimierte Variante führt bei gleichzeitiger Expression des pflanzlichen

GGPPS-Gens zu keiner erhöhten Taxadienproduktion, sondern sogar zu einer leichten Verringerung von 0,31 mg/l auf 0,23 mg/l. Die Substitution des pflanzlichen GGPPS-Gens durch das archaebakterielle Gen führte dagegen zu einer 100 fachen Steigerung des detektierbaren GGOHs von 0,29 mg/l auf 33,1 mg/l in Kombination mit dem synthetischen Taxadiensynthesegen, bzw. von 0,28 mg/l auf 27,6 mg/l bei Koexpression mit dem pflanzlichen *TDS*. Wird das codon-optimierte TDS-Gen allerdings mit dem archaebakteriellen *GGPPS* kombiniert, steigt die gemessene Menge an Taxadien 40-fach von 0,23 mg/l auf 8,7 mg/l. Abbildung III.7 zeigt die Chromatogramme der GC/MS-Untersuchungen von CEN 4-7.



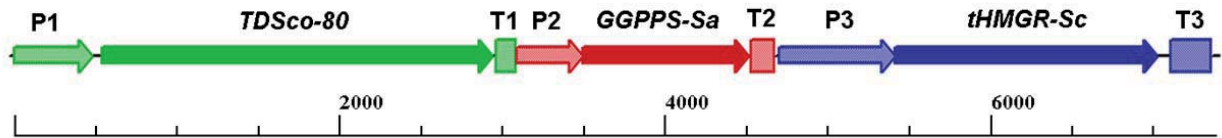
**Abbildung III.7: Gaschromatographische Auftrennung der Extrakte Taxadien produzierender Hefestämme.** Die Identifizierung der jeweiligen Substanzen erfolgte anhand ihrer Massenspektren. **A:** CEN 4 (pVV214::*TDS-Tc*; pVV200::*GGPPS-Tc*; pRS315::*tHMGR-Sc*; pRS313::*UPC2.1-Sc*); **B:** CEN 5 (pVV214::*TDS-Tc*; pVV200::*GGPPS-Sa*; pRS315::*tHMGR-Sc*; pRS313::*UPC2.1-Sc*); **C:** CEN 6 (pVV214::*TDSco*; pVV200::*GGPPS-Tc*; pRS315::*tHMGR-Sc*; pRS313::*UPC2.1-Sc*); **D:** CEN 7 (pVV214::*TDSco*; pVV200::*GGPPS-Sa*; pRS315::*tHMGR-Sc*; pRS313::*UPC2.1-Sc*); **TD:** Taxadien; **IS:** Interner Standard 4,4'-di-tert-butylbiphenyl; **GGOH:** Geranylgeraniol

### III.3 Generierung eines Genclusters

Der generierte Hefestamm CEN 7 (III.1) ist nachweislich in der Lage, die bisher größte bekannte Menge an Taxadien zu produzieren. Allerdings besteht der Taxolbiosyntheseweg aus ca. 19 enzymkatalysierten Reaktionen, von denen die Zyklisierung von GGPP zu Taxadien nur den einleitenden Schritt darstellt (Abbildung I.16). Die Biosynthese späterer Intermediate erfordert daher die heterologe Expression zusätzlicher Taxolbiosynthesegene. Die Auxotrophien Histidin, Leucin, Uracil und Tryptophan in CEN 7 sind durch die eingebrachten Plasmide bereits komplementiert, die Selektion auf weitere Vektoren mit zusätzlichen Genen in diesem Stamm folglich ohne Verwendung dominanter Selektionsmarker nicht durchführbar. Eine Fusionierung mehrerer Gene würde ihre Coexpression auf einem gemeinsamen Vektor zulassen und somit die Selektion dieser Gene über den gleichen Marker ermöglichen. Die Anzahl an Auxotrophien des *S. cerevisiae*-Stamms CEN-PK2-1C wäre demnach für die Expression weiterer Gene nicht mehr limitierend.

#### III.3.1 Klonierung

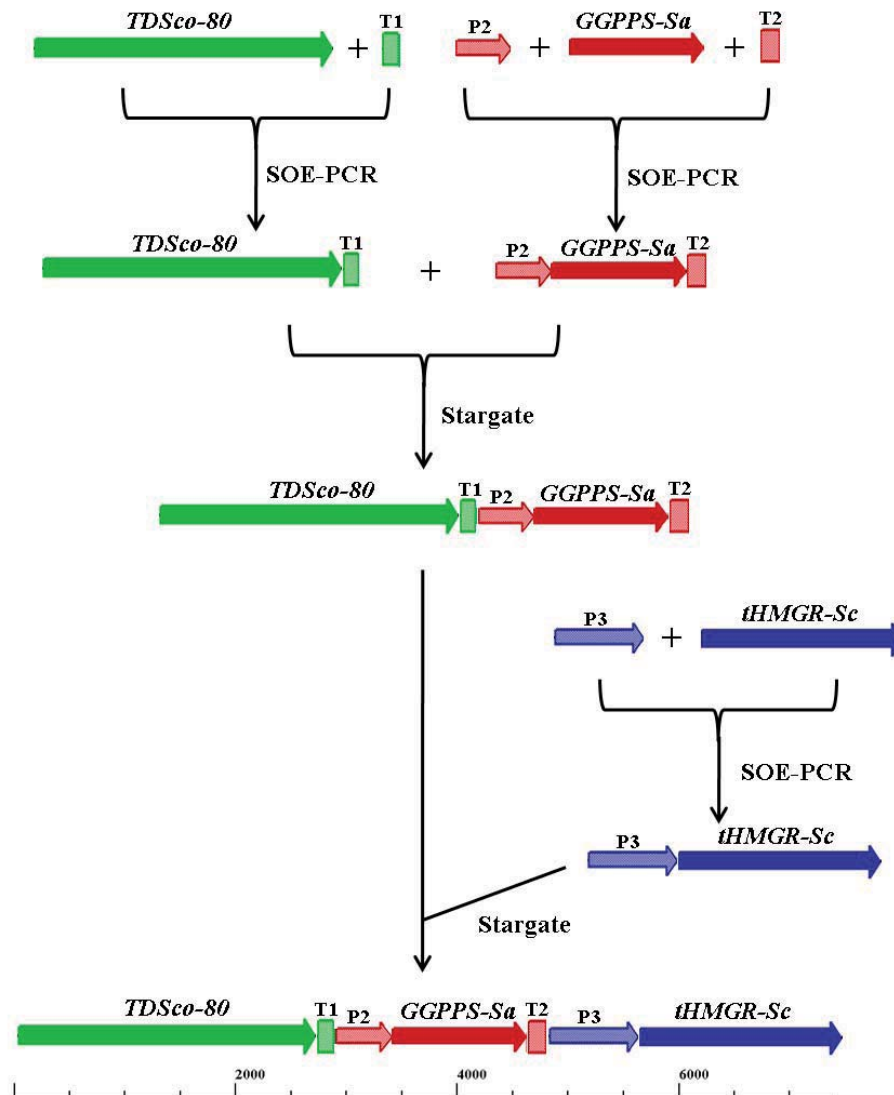
Basierend auf der quantitativen Produktanalyse mittels GC/MS (Tabelle III.3) wurden die Auswirkungen einer Expression von *UPC2.1-Sc* als gering eingeschätzt, sodass die Fusionierung auf die Gene *TDSco*, *GGPPS-Sa* und *tHMGR-Sc* beschränkt wurde. Diese sollten mit jeweils eigenem Promotor und Terminator in dem in Abbildung III.8 dargestellten Konstrukt zusammengeführt werden. Zeitgleich wurden im Rahmen einer Diplomarbeit Untersuchungen an *TDSco* verschiedener Länge durchgeführt. Das bisher verwendete, für die *N*-terminal um 60 Aminosäuren verkürzte *TDSco* kodierende Gen wurde dazu durch solche ersetzt, die für um 64, 71 und 80 Aminosäuren verkürzte *TDSco* kodieren. Anhand von Westernblots in jener Arbeit konnte gezeigt werden, dass die -80-Variante (*TDSco*-80) die stärkste Proteinakkumulation aufweist. GC/MS-Messungen ergaben für den *TDSco*-80 exprimierenden Stamm des Weiteren die erfolgreichste Taxadien-Produktionsleistung. Aufgrund dieser Ergebnisse wurde bei der Generierung des Clusters das für eine um 80 Aminosäuren verkürzte *TDSco* kodierende Gen verwendet.



**Abbildung III.8: Geplantes Fusionskonstrukt.** **P1:** *GPD*-Promotor; ***TDSco-80*:** codon-optimiertes Gen für Taxadiensynthase, *N*-terminal um 80AS verkürzt; **T1:** *PHO5*-Terminator; **P2:** *GAP1*-Promotor; ***GGPPS-Sa*:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **T2:** *GCN4*-Terminator; **P3:** *ADH1*-Promotor; ***tHMGR-Sc*:** Gensequenz für die katalytische Domäne der HMG CoA-Reduktase Isozym 1 aus *S. cerevisiae*; **T3:** *CYC1*-Terminator

Die Fusionierung der drei Gene mit ihren jeweiligen Regulationseinheiten erfolgte anhand von SOE-PCR (II.2.2.2). Die anschließende Verknüpfung der Einheiten wurde dagegen mithilfe der Stargate-Technologie durchgeführt (II.2.6.2). Als Grundgerüst des geplanten Fusionsvektors diente der Gateway-kompatible Vektor pAG423 (Tabelle II.2), der sowohl den Promotor für das innerhalb des Konstrukts 5'-gelegene *TDSco*, als auch den Terminator für das 3'-lokalisierte *tHMGR-Sc* bereitstellt. Über SOE-PCR wurde daher *TDSco-80* ausschließlich mit dem Phosphatase-Terminator *pho5tt* (Primer 15-18) und *tHMGR-Sc* mit dem Alkoholdehydrogenase-Promotor *adh1p* verbunden (Primer 17-20). An das mittlere Gen *GGPPS-Sa* wurde sowohl der Glycerinaldehyd-3-Phosphat-Dehydrogenase-Promotor *gap1p* (Primer 21-24), als auch der Terminator des Transkriptionsaktivatorproteins *gcn4tt* (Primer 23-26) gehängt. Die Regulationseinheiten wurden aus genomischer Hefe-DNS amplifiziert, während den Genen der jeweilige pVV-Vektor als Matrize diente (Tabelle III.1). Die Sequenzen der verwendeten Primer sind in Tabelle II.3 angegeben. Die SOE-PCR Produkte „*TDSco80-pho5tt*“, „*gap1p-GGPPS-Sa-gcn4tt*“ und „*adh1p-tHMGR-Sc*“ wurden anschließend mithilfe der Stargate-Technologie (II.2.6.2) schrittweise fusioniert. Der Einsatz der jeweiligen Methode bei der Fusionierung ist in Abbildung III.9 zusammengefasst.





**Abbildung III.9: Fusionierungsansatz der Gene *TDSc0-80* (grün), *GGPPS-Sa* (rot) und *tHMGR-Sc* (blau).** Nach Verlinkung der Gene mit Promotor- (kleine Pfeile) und Terminatorsequenz (Kästchen) über SOE-PCR, folgte die schrittweise Fusionierung dieser Einheiten via Stargate. T1: *PHO5*-Terminator; P1: *GAP1*-Promotor; T2: *GCN4*-Terminator; P3: *ADH1*-Promotor

Die für die Stargate-Fusionierung benötigten Rekombinationssequenzen wurden im Rahmen einer PCR durch die verwendeten Primer 27-32 angehängt (Tabelle II.3). Um eine anschließende Rekombination der fusionierten DNS in Gateway-kompatible Expressionsvektoren zu ermöglichen, wurden den endständigen Fragmenten des späteren Fusionskonstrukts *TDSc0-80* und *tHMGR-Sc* zusammen mit den Stargate-Rekombinationssequenzen auch Gatewayeinheiten angehängt (Abbildung II.6).

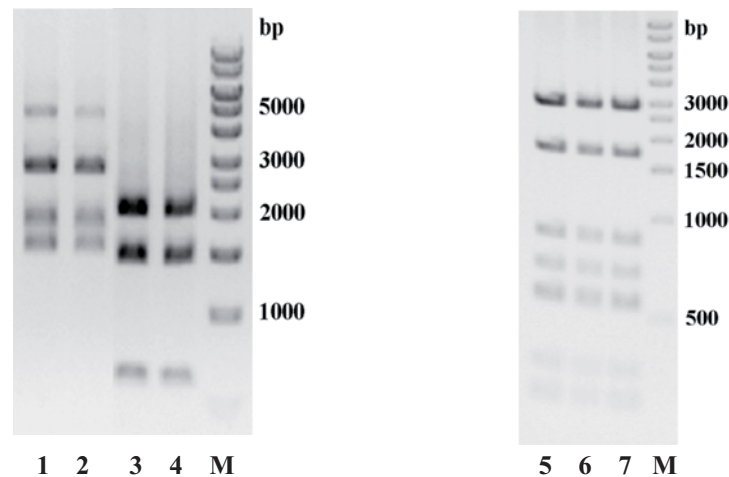


Alle im Folgenden beschriebenen Stargate-Vektoren sind in Tabelle III.5 zusammengefasst. Für die Verbindung mehrerer DNS-Abschnitte via Stargate-Technologie müssen diese vom generierten Donorvektor in einen Fusionsvektor rekombiniert werden. Abhängig von der gewünschten Position innerhalb des Endkonstrukts wird dafür ein *upstream* - Fusionsvektor (für die 5'-Position) oder ein *downstream*-Fusionsvektor (für die spätere 3'-Position) verwendet. Der durch die initiale Stargate-Reaktion generierte pDONOR::*TDSc80-pho5tt* (S1) wurde daher weiter zum *upstream*-Fusionsvektor pNFUSE::*TDSc80-pho5tt* (S4) rekombiniert. Um den DNS-Abschnitt *gap1p-GGPPS-Sa-gcn4tt* damit zu verbinden, wurde erst pDONOR::*gap1p-GGPPS-Sa-gcn4tt* (S2) generiert und dieser zu dem *downstream*-Fusionsvektor pCFUSE::*gap1p-GGPPS-Sa-gcn4tt* (S5) rekombiniert. Die Fusionsvektoren S4 und S5 ergaben anschließend das Fusionsprodukt pDONOR::*TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt* (S6). Für die weitere Verlinkung mit *adh1p-tHMGR-Sc* wurde erst pDONOR::*adh1p-tHMGR-Sc* (S3) und daraus pCFUSE::*adh1p-tHMGR-Sc* (S8) generiert. Der *downstream*-Fusionsvektor S8 wurde mit dem *upstream* - Fusionsvektor pNFUSE::*TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt* (S7) abschließend zum finalen Fusionsprodukt pDONOR::*TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc* (S9) rekombiniert.

**Tabelle III.5: Generierte Stargate-Vektoren**

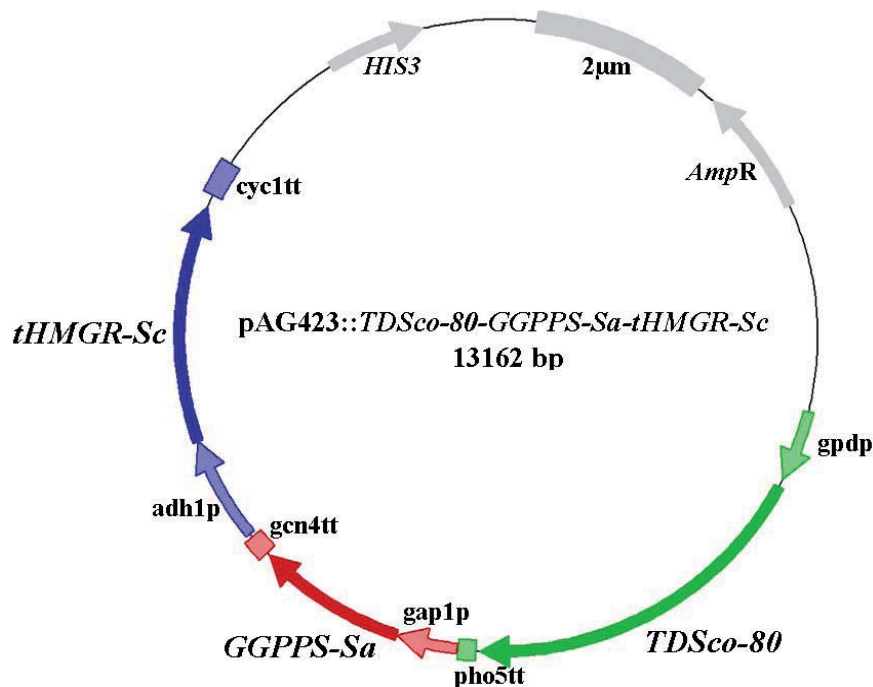
Nr	Plasmid	Funktion
S1	pDONOR:: <i>TDSc80-pho5tt</i>	Eingangsvektor für die Rekombination zum 5'- oder 3'-Fusionsvektor
S2	pDONOR:: <i>gap1p-GGPPS-Sa-gcn4tt</i>	Eingangsvektor für die Rekombination zum 5'- oder 3'-Fusionsvektor
S3	pDONOR:: <i>adh1p-tHMGR-Sc</i>	Eingangsvektor für die Rekombination zum 5'- oder 3'-Fusionsvektor
S4	pNFUSE:: <i>TDSc80-pho5tt</i>	5'-Fusionsvektor zur Rekombination mit 3'-Fusionsvektor S5
S5	pCFUSE:: <i>gap1p-GGPPS-Sa-gcn4tt</i>	3'-Fusionsvektor zur Rekombination mit 5'-Fusionsvektor S4
S6	pDONOR:: <i>TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt</i>	Fusionsprodukt von S4 mit S5
S7	pNFUSE:: <i>TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt</i>	5'-Fusionsvektor zur Rekombination mit 3'-Fusionsvektor S8
S8	pCFUSE:: <i>adh1p-tHMGR-Sc</i>	3'-Fusionsvektor zur Rekombination mit 5'-Fusionsvektor S7
S9	pDONOR:: <i>TDSc80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc</i>	Fusionsprodukt von S7 mit S8

Alle generierten Stargate-Vektoren wurden mittels Restriktion (II.2.5.1) auf erfolgreiche Rekombinationsereignisse getestet. Eine gelelektrophoretische Auftrennung der durchgeführten Restriktionsanalysen ergab für alle getesteten Fusionskonstrukte das jeweils korrekte Bandenmuster. In Abbildung III.10 sind exemplarisch die Restriktionsansätze der in den letzten Reaktionsschritt eingesetzten Fusionsvektoren sowie des Endprodukts dargestellt.



**Abbildung III.10: Restriktionsanalyse der im letzten Fusionierungsschritt involvierten Stargate-Vektoren. 1-2:** *Bci*VI-Verdau von *upstream*-Fusionsvektoren pNFUSE::*TDSco80*-pho5tt-gap1p-*GGPPS*-*Sa*-gcn4tt. Erwartetes Restriktionsmuster: 2,7kb + 1,7kb + 1,5kb **3-4:** *Bci*VI-Verdau von *downstream*-Fusionsvektoren pCFUSE::*adh1p-tHMGR-Sc*. Erwartetes Restriktionsmuster: 2,1kb + 1,5kb + 620bp; **M:** NEB-Größenstandard; **5-7:** *Xba*I/*Hind*III-Verdau finaler Fusionsvektoren pDONOR::*TDSco80*-pho5tt-gap1p-*GGPPS*-*Sa*-gcn4tt-*adh1p-tHMGR-Sc*. Erwartetes Restriktionsmuster: 2,9kb + 1,7kb + 880bp + 700bp + 550bp + 360bp + 250bp + 60bp.

Das Endprodukt der Stargate-Fusionierung (S9) wurde mittels Gateway-Technologie in den Vektor pAG423 (Tabelle II.2) rekombiniert. Ermöglicht wurde dies durch die im Rahmen der Stargate-PCR integrierten Gateway-Sequenzen (Abbildung II.6). Abbildung III.11 zeigt die Vektorkarte des generierten Fusionskonstrukts.



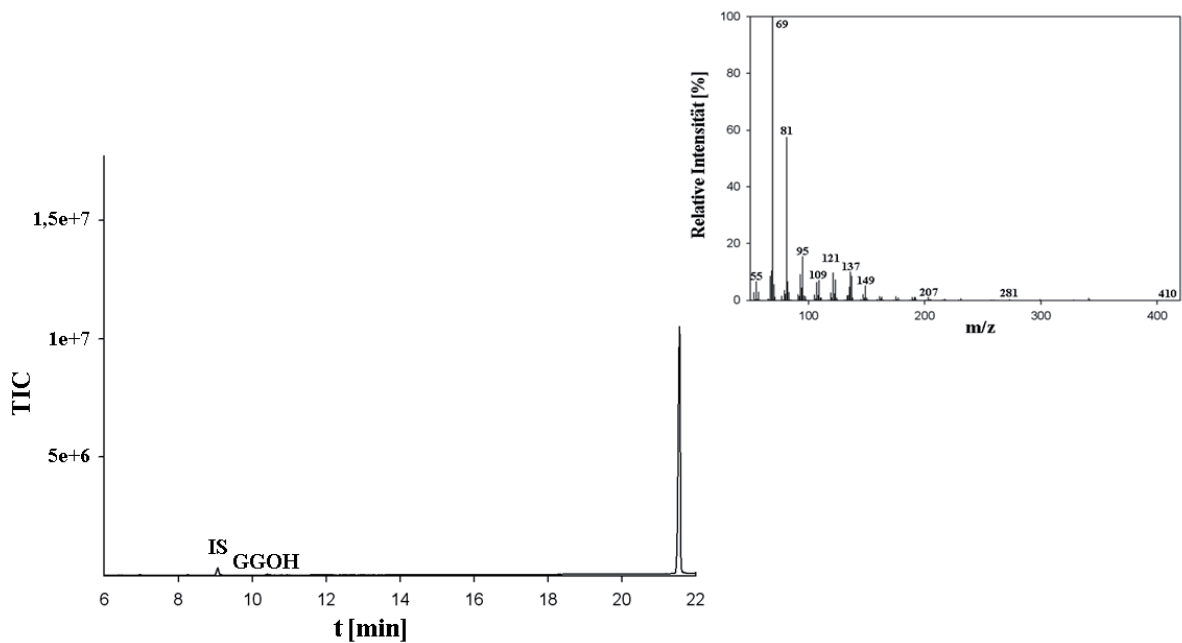
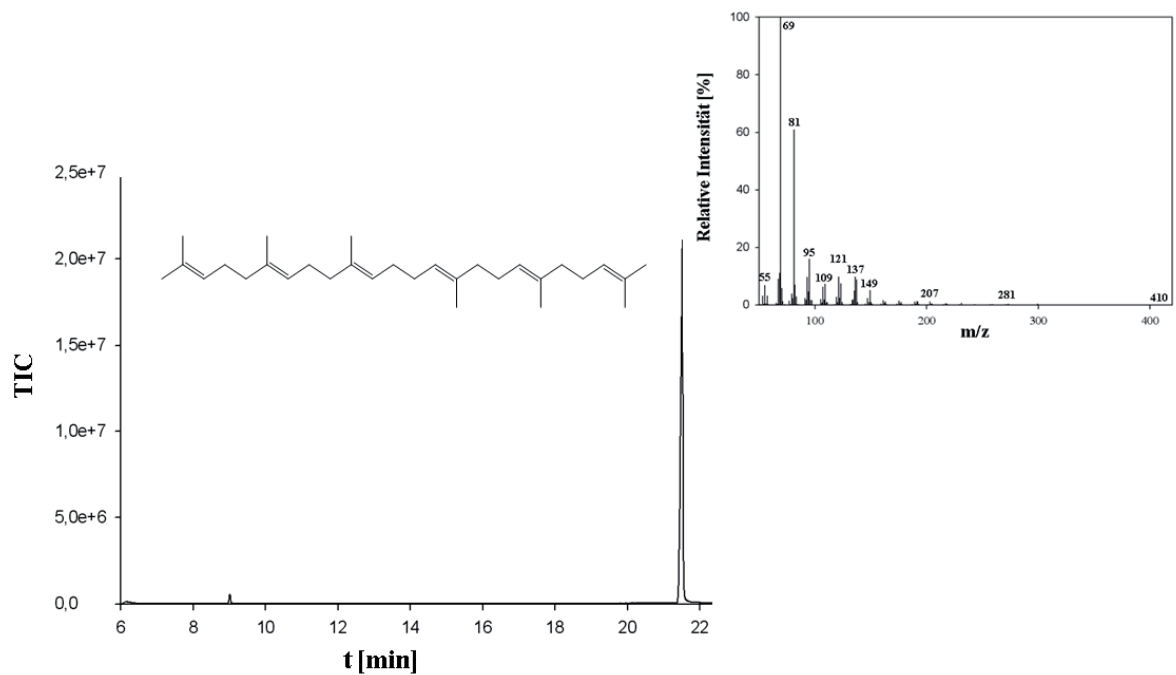
**Abbildung III.11: Vektorkarte des generierten Fusionskonstrukts.** *HIS3*: Selektionsmarker Histidin; *2µm*: 2micron; *AmpR*: Ampicillin-Resistenzmarker; *gdp*: Glukose-6-phosphat Dehydrogenase Promotor; *TDSco-80*: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; *pho5tt*: Saure Phosphatase-Terminator; *gap1p*: Glycerinaldehyd-3-Phosphat-dehydrogenase-Promotor; *GGPPS-Sa*: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; *gcn4tt*: Terminator des Gens für Transkriptionsaktivatorprotein; *adh1p*: Alkoholdehydrogenase-Promotor; *tHMGR-Sc*: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; *cyc1tt*: Cytochrom1-Terminator

Die korrekte Basenabfolge des Endkonstrukts pAG423::TDSco80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc wurde durch Sequenzierung bestätigt. Dafür verwendete Oligonukleotidprimer sind in Tabelle II.4 aufgelistet.

### III.3.2 Stammanalyse

Mit dem generierten Fusionsvektor pAG423::TDSco80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc (pAG423::T-G-t) wurde der *S. cerevisiae*-Stamm CEN.PK2-1C (Tabelle II.1) transformiert und der entstandene Stamm CEN 8 mittels GC/MS auf die Synthese von Taxadien und GGOH untersucht. Dabei konnte eine Taxadienkonzentration von 0,41 mg/l, sowie eine GGOH-Produktion von 0,34 mg/l (Tabelle III.6) detektiert werden. Dies entspricht einem Zwanzigstel der Taxadien- und einem Hundertstel der GGOH-Produktion verglichen mit den auf Einzelvektoren exprimierten Genen. Bei Betrachtung des Chromatogramms fiel ein Signal mit einer

Retentionszeit von 21,5 min auf, das mit einem Basispeak von 69 ein für lineare Terpene typisches Massenspektrum aufwies. Aufgrund seiner fortgeschrittenen Retentionszeit wurde vermutet, dass es sich um das Triterpen Squalen handeln könnte. GC/MS-Messungen kommerziellen Squalens bestätigten diese Hypothese (Abbildung III.12).



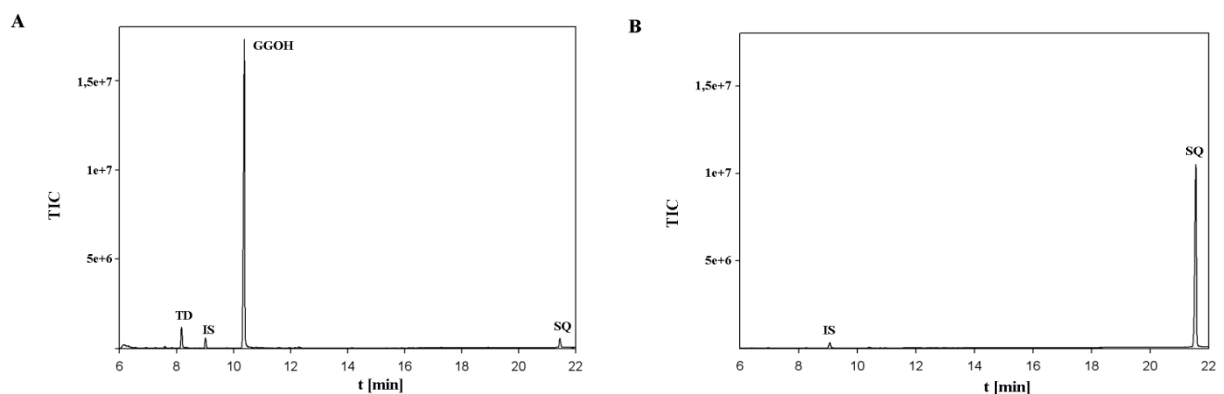
**Abbildung III.12: GC/MS-Analyse kommerziellen Squalens (A) im Vergleich mit Extrakten aus CEN 8.** Sowohl die Retentionszeit von 21,5 min als auch das Massenspektrum [69(100%), 81(59%), 95(15%), 121(9%), 137(8%), 55(7%) mit einem Molekülion von 410] bestätigen die Akkumulation von Squalen im Fusionsstamm.

Anhand einer Standardgeraden (II.3.5.1) konnte eine Squalenproduktion des Stamms CEN 8 von 86,3 mg/l bestimmt werden (Tabelle III.6).

**Tabelle III.6 : Vergleich gebildeter Produktmengen via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±5,6	3,90±0,3
CEN 8	<i>TDSco-80-GGPPS-Sa-tHMGR1-Sc</i>	0,41±0,02	0,34±0,05	86,3±12

Ein Vergleich der Squalenkonzentration des Fusionsstamms mit denjenigen des Wildtyps und CEN 7 ergab eine 28 fache Steigerung im Fusionsstamm gegenüber dem Wildtyp-Stamm CEN.PK2-1C und einen Anstieg auf das 22 fache verglichen mit dem die Einzelgene exprimierenden Stamm CEN 7 (Tabelle III.6). Abbildung III.13 gibt die Chromatogramme der mittels GC/MS auf ihre Produktion untersuchten Stämme CEN 7 und CEN 8 wieder. Aufgetragen ist jeweils die absolute Intensität gegen die Retentionszeit.



**Abbildung III.13: Gaschromatographische Auftrennung der Extrakte generierter Hefestämme. A:** CEN (pVV214::TDSco; pVV200::GGPPS-Sa; pRS315::tHMGR-Sc; pRS313::UPC2.1-Sc); **B:** CEN (pAG423::TDSco-80-GGPPSSa-tHMGRSc); **TD:** Taxadien; **IS:** Interner Standard 4,4-di-tert-butylbiphenyl; **GGOH:** Geranylgeraniol; **SQ:** Squalen.

Um das Auftreten der geringen Menge an gebildetem Taxadien und GGOH zu analysieren, wurde der Fusionsstamm durch Westernblotanalyse auf die Synthese der Zielproteine hin untersucht (Abbildung III.14). Die Negativkontrolle bildete dabei der Wildtyp-Hefestamm. Um eine vergleichbare Proteinmenge der untersuchten Stämme zu garantieren, wurde parallel eine SDS-PAGE durchgeführt und die aufgetrennten Proteine mit Coomassieblau visualisiert (II.3.2).

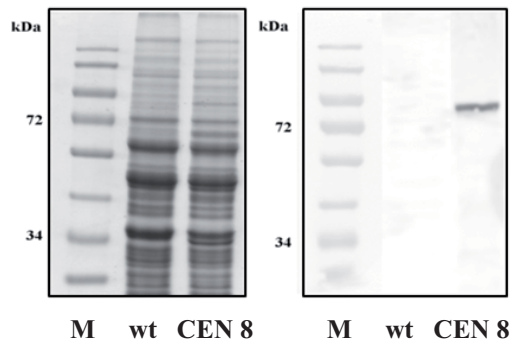
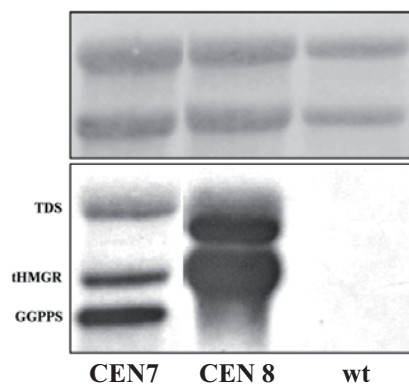


Abbildung III.14: SDS-PAGE (links) und Western blot (rechts) von 40 µg TCA-gefälltem Gesamtprotein des Wildtyps sowie des generierten Fusionsstamms. M: **Größenstandard (NEB)**; wt: **CEN.PK2-1C Wildtyp**; CEN 8: **CEN.PK2-1C (pAG423::TDSco-80-GGPPSSa-tHMGR)**

Die Westernblotanalyse des Stamms CEN 8 ergab zwar eine Akkumulation von TDSco, aber weder von GGPPS-Sa noch von tHMGR-Sc (Abbildung III.14). Der Fusionsstamm wurde daher einem Northernblot unterzogen, um ihn auf die Expression von *GGPPS-Sa* und *tHMGR-Sc* zu untersuchen. Abbildung III.15 zeigt die Signale der spezifisch gebundenen radioaktiv markierten DNS-Sonde der im Fusionsstamm exprimierten Gene. Als Positivkontrolle wurde parallel CEN 7 (Tabelle III.2) und als Negativkontrolle der Wildtypstamm mitgeführt. Die Sequenzen der als Sonde eingesetzten Oligonukleotide sind in Tabelle II.5 angegeben. Trotz fehlenden Signals auf Proteinebene konnte *tHMGR-Sc* auf mRNA-Ebene detektiert werden. Ein Nachweis von *GGPPS-Sa*-mRNA konnte dagegen nicht erbracht werden.



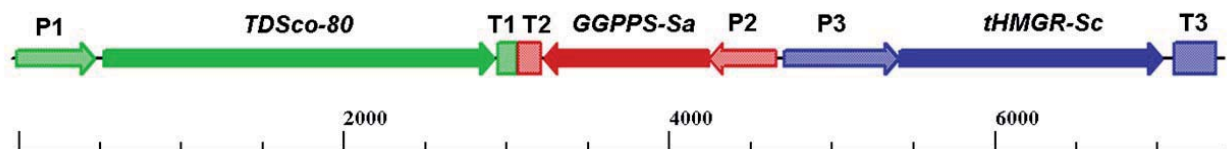
**Abbildung III.15: Northernblotanalyse des generierten Fusionsstamms. Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10µg geladen. **Unten:** spezifisch gebundene radioaktiv markierte DNS, auf Röntgenfilm visualisiert; **CEN 7:** CEN (pVV214::TDSco; pVV200::GGPPS-Sa; pRS315::tHMGR-Sc; pRS313::UPC2.1-Sc); **CEN 8:** CEN (pAG423::TDSco-80-GGPPSSa-tHMGRSc); **Wt:** CENPK2-1C Wildtyp

### III.3.3 Modifizierung des Genclusters

Die fehlende Expression von *GGPPS-Sa* (Abbildung III.15) ist durch die Wahl des Promotors nicht zu erklären. Beide internen Promotorsequenzen des Fusionskonstrukts (*GAP1*-Promotor vor *GGPPS-Sa*, bzw. *ADH1*-Promotor vor *tHMGR-Sc*) repräsentieren ursprünglich Regulationselemente für Gene, deren Produkte in die Glykolyse involviert sind. Aufgrund konstitutiver Expression ihrer Gene liegen diese Enzyme in hoher Konzentration vor, was im Falle der ubiquitären Glycerinaldehyd-3-Phosphat-Dehydrogenase (GAP) in Anteilen von bis zu 5 % der Zelltrockenmasse von Bäckerhefen resultiert. Die jeweiligen Sequenzen wurden aus isolierter genomischer *S. cerevisiae*-DNS amplifiziert. Eine genaue Übereinstimmung der in das Fusionskonstrukt eingesetzten Promotoren mit der genomischen Basenabfolge wurde durch Sequenzierungen bestätigt. Allerdings wurden die einzelnen Transkriptionseinheiten ohne Spacer zwischen Terminator und Promotor aneinander grenzender Gene fusioniert, sodass die Bildung eines Promotorkomplexes während der Transkription des einen Gens diejenige eines benachbarten Gens behindern könnte.<sup>[213]</sup> Dagegen spricht, dass das *Saccharomyces cerevisiae*-Genom selbst sehr kompakt ist. Über 72 % der Sequenzen sind in ORFs organisiert, sodass nicht-kodierende Sequenzen wie Promotoren, Terminatoren und andere regulatorische Elemente einen relativ kleinen Teil ausmachen. Basierend auf Berechnungen von Dujon konnte anhand der Distanzen benachbarter, divergenter ORFs so eine durchschnittliche Promotorlänge von 386 bp, anhand der Distanzen benachbarter konvergenter ORFs eine Terminatorlänge von durchschnittlich 145 bp bestimmt werden.<sup>[214]</sup> Diese Angaben werden von den im Rahmen dieser Arbeit selektierten Promotorsequenzen von 404 bp (*gap1p*) und 703 bp (*adh1p*) deutlich übertroffen, sodass zu kurz gewählte Regulationseinheiten ausgeschlossen werden können. Auch die gewählte Orientierung der fusionierten Gene als Tandemkonstrukt entspricht der in *Saccharomyces cerevisiae* bevorzugten Organisation. Nach Beendigung des *Yeast genome sequencing program* 1996 konnte die Anordnung dort auf 50 % an Tandemgenen und jeweils 25 % an konvergent, bzw. divergent orientierten Genen festgestellt werden.<sup>[215]</sup> Untersuchungen an den benachbarten Genen *ARO4* und *HIS7*, die wie das konstruierte Gencluster im Tandem organisiert sind, zeigten des Weiteren eine aktive Expression beider Gene trotz fehlender intergener Region.<sup>[216][217]</sup> Am Beispiel der gleichzeitig exprimierten Gene *GAL10* und *GAL7* wurde dennoch demonstriert, dass die Expression des 3'-gelegenen Gens durch die Expression des 5'-lokalisierten Gens inhibiert werden kann, wenn die Transkription des ersten Gens initiiert, aber nicht korrekt terminiert wird. Dabei verdeckt die RNS-Polymerase II durch Transkription des ersten Gens den Promotor des folgenden Gens und verhindert so die Bindung der benötigten Transkriptionsfaktoren.<sup>[218]</sup> Die beschriebene



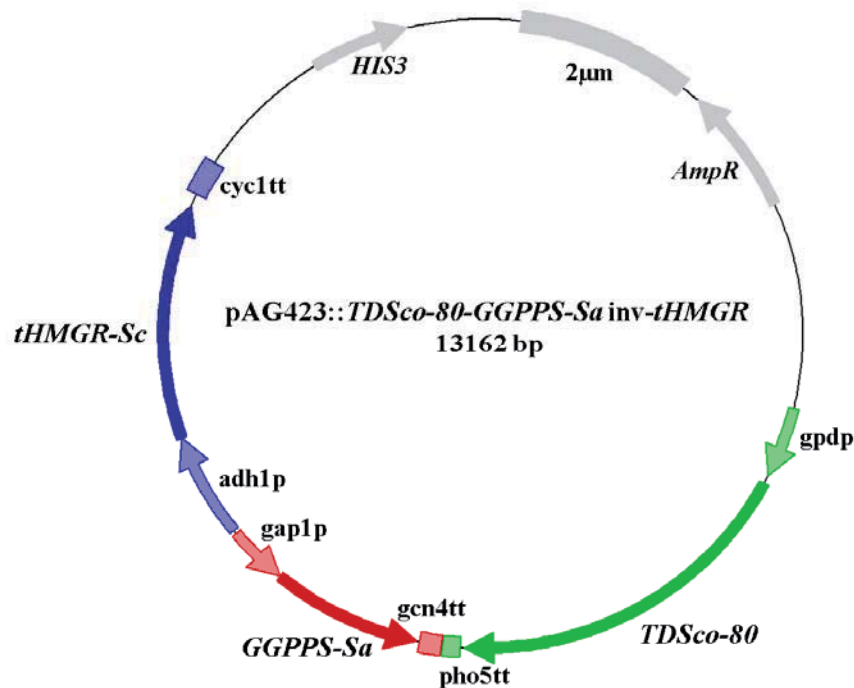
transkriptionale Interferenz wird dabei auf eine nicht terminierte Aktivität der Polymerase zurückgeführt, was zur Bildung von bicistronischer mRNA verschiedener Länge führen würde. Anhand einer Northernblotanalyse konnte im Rahmen dieser Arbeit jedoch eine definierte mRNA-Länge des transkribierten *TDSco-80* nachgewiesen werden. Allerdings wurden transkriptionale Interferenzen (TI) auch unabhängig von einer Termination der transkribierenden Polymerase beobachtet. Ein weiterer vorgeschlagener Mechanismus basiert daher auf der rein sterischen Behinderung von Transkriptionskomplexen zweier benachbarter Gene.<sup>[213]</sup> Im Falle des Fusionskonstrukts würde das erste Gen *TDSco* transkribiert mit der Folge einer inhibierten Expression von *GGPPS-Sa*. Die unterbundene Bindung der RNS-Polymerase an *GGPPS-Sa* wiederum würde die problemlose Transkription von *tHMGR* ermöglichen. Obgleich sowohl bei konvergent als auch divergent orientierten benachbarten Genen eine Beeinträchtigung der Expression nachgewiesen wurde,<sup>[219][220]</sup> konnte gezeigt werden, dass die Anordnung divergenter ORFs die geringste Auswirkung auf die Expression beider Gene hat.<sup>[213]</sup> Daher wurde das ursprüngliche Gencluster *TDSco-80-GGPPS-Sa-tHMGR-Sc* so modifiziert, dass eine Tandemanordnung zweier Gene vermieden wird und das bisher nicht exprimierte Gen *GGPPS-Sa* in divergenter Orientierung zu *tHMGR-Sc* steht (Abbildung III.16).



**Abbildung III.16: modifiziertes Fusionskonstrukt.** Die Gene *TDSco-80* und *GGPPS-Sa* sind konvergent orientiert, *GGPPS-Sa* und *tHMGR-Sc* stehen divergent zueinander. **P1:** GPD-Promotor; ***TDSco-80*:** codon-optimiertes Gen für Taxadien-synthase, N-terminal um 80 AS verkürzt; **T1:** *PHO5*-Terminator; **P2:** *GAP1*-Promotor; ***GGPPS-Sa*:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **T2:** *GCN4*-Terminator; **P3:** *ADH1*-Promotor; ***tHMGR-Sc*:** Gensequenz für die katalytische Domäne der HMGCoA-Reduktase Isozym 1 aus *S. cerevisiae*; **T3:** *CYC1*-Terminator

Um das DNS-Fragment *gap1p-GGPPS-Sa-gcn4tt* umzupositionieren, wurde es mit phosphorylierten Primern amplifiziert (Primer 32 und 34, Tabelle II.3) und anschließend mit dem ebenfalls amplifizierten Vektorfragment *pho5tt-TDSco-80-pAG423-tHMGR-Sc-adh1p* ligiert (Primer 16 und 17, Tabelle II.3). Der generierte Fusionsvektor *pAG423::TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-gap1p-adh1p-tHMGR-Sc* (*pAG423-T-G-t inv*) ist in Abbildung III.17 dargestellt.





**Abbildung III.17: Vektorkarte des Fusionskonstrukts pAG423::T-G-t inv.** Die Gene *TDSco-80* und *GGPPS-Sa* sind konvergent orientiert, *GGPPS-Sa* und *tHMGR-Sc* stehen divergent zueinander. **HIS3**: Selektionsmarker Histidin; **2µm**: 2micron Replikationsursprung; **AmpR**: Ampicillin-Resistenzmarker; **gdp**: Glukose-6-phosphat Dehydrogenase-Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **pho5tt**: Saure Phosphatase-Terminator; **gap1p**: Glycerinaldehyd-3-Phosphat-dehydrogenase-Promotor; **GGPPS-Sa**: Geranylgeranyl-diphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **gcn4tt**: Terminator des Gens für Transkriptionsaktivatorprotein; **adh1p**: Alkoholdehydrogenase-Promotor; **tHMGR-Sc**: verkürztes HMGR-Gen aus *S. cerevisiae*; **cyc1tt**: Cytochrom 1-Terminator

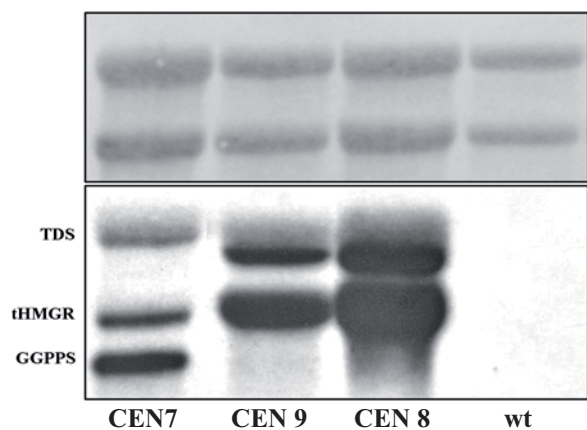
Nach Bestätigung der richtigen Basenabfolge durch Sequenzierung wurde der modifizierte Fusionsvektor verwendet, um den *S. cerevisiae*-Stamm CEN.PK2-1C zu transformieren und der generierte Stamm CEN 9 einer Produktanalyse mittels GC/MS unterzogen. Tabelle III.7 gibt die erhaltenen Konzentrationen an extrahiertem Taxadien, GGOH und Squalen wieder.

**Tabelle III.7 : Quantifizierung gebildeter Produktmengen via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±5,6	3,90±0,3
CEN 8	<i>TDSco-80-GGPPS-Sa-tHMGR1-Sc</i>	0,41±0,02	0,34±0,05	86,3±12
CEN 9	<i>TDSco-80-GGPPS-Sa inv-tHMGR1-Sc</i>	0,42±0,005	0,45±0,03	94,8±21

Die Umordnung der fusionierten Gene führte mit 0,42 mg/l zu keiner Veränderung des Taxadiengehalts. Dagegen konnte sowohl ein Anstieg der GGOH-Produktion von 0,34 mg/l auf 0,45 mg/l, als auch eine Steigerung des Squalengehalts um 8,5 mg/l auf 94,80 mg/l detektiert werden.

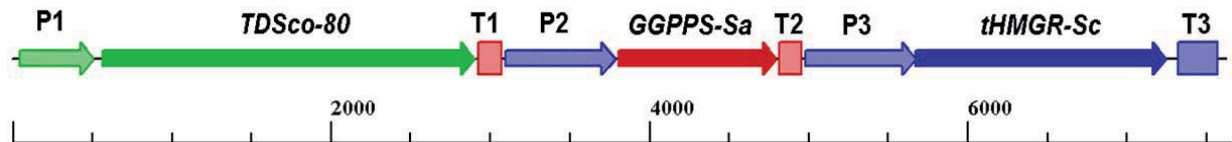
Trotz nachweisbarer mRNA von *TDSco-80* und *tHMGR-Sc*, konnte eine Transkription des invertierten *GGPPS-Sa* durch einen Northernblot nicht bestätigt werden (Abbildung III.18). Die durchgeführte Westernblotanalyse ergab weiterhin keine nachweisbare Translation möglicherweise synthetisierter mRNA.



**Abbildung III.18: Northernblotanalyse generierter Fusionsstämme. Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10 µg geladen. **Unten:** Autoradiogramm der spezifisch gebundenen radioaktiv markierten DNS-Sonde; **CEN 7:** Einzeln exprimierte Gene *TDSco-*, *GGPPS-Sa* und *tHMGRSc*; **CEN 8:** *TDSco-80*, *GGPPS-Sa* und *tHMGRSc* als Tandemkonstrukt angeordnet; **CEN 9:** *TDSco-80* und *GGPPS-Sa* konvergent, *GGPPS-Sa* und *tHMGRSc* divergent orientiert; **Wt:** CENPK2-1C Wildtyp

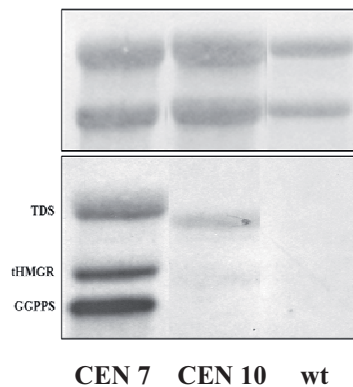
Sowohl das erste (pAG423::T-G-t), als auch das zweite Fusionskonstrukt (pAG423::T-G-t inv) zeigten eine Expression der beiden randständigen Gene *TDSco-80* und *tHMGR-Sc*. Das mittlere Gen dagegen wird nicht transkribiert (Abbildung III.18), obwohl der identische ORF im Einzelvektor (pVV200::*GGPPS-Sa*) exprimiert in einem mRNA- und Proteinsignal resultiert (Abbildung III.3, Abbildung III.4). Die Orientierung der einzelnen Einheiten zueinander ist für die Bildung von *GGPPS-Sa*-mRNA offenbar nicht ausschlaggebend. Alle Gene wurden ohne Spacer mit ihrem Promotor und Terminator fusioniert und diese Einheiten anschließend über Stargate-Technologie verbunden. Der Abstand zwischen Gen und Regulationseinheiten, bzw. benachbarter Regulationseinheiten ist also identisch. Was die einzelnen Abschnitte unterscheidet, ist die Wahl des Promotors und Terminators. *TDSco-80* ist der vektoreigene *GPD*-Promotor vorgeschaltet. Die Transkription von *tHMGR-Sc* wird durch den in pAG423 integrierten *PHO5*-Terminator beendet. Von den clusterinternen Regulationseinheiten führt ausschließlich die zwischen *GGPPS-Sa* und *tHMGR-Sc* geschaltete Terminator-Promotor-Einheit *gcn4tt-adh1p* zu einer aktiven Expression. Um

eine Transkription von *GGPPS-Sa* zu erhalten, wurde daher das scheinbar funktionelle Fragment *gcn4tt-adh1p* auch zwischen *TDSCO-80* und *GGPPS-Sa* positioniert (Abbildung III.19).



**Abbildung III.19: Fusionskonstrukt T-G-t GA.** Die Gene *TDSCO-80*, *GGPPS-Sa* und *tHMGR-Sc* sind in gleicher Orientierung hintereinander geschaltet. Die Regulationseinheiten zwischen den Genen sind identisch (*gcn4tt-adh1p*). **P1:** *GPD*-Promotor; **T1 und T2:** *GCN4*-Terminator; **P2 und P3:** *ADH1*-Promotor; **T3:** *CYC1*-Terminator

Die Modifizierung des Fusionsvektors erfolgte erneut mit Stargate-Technologie. Sowohl *TDSCO-80*, als auch *gcn4tt-adh1p* und *GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc* wurden dafür erst in einen pDONOR und anschließend in Fusionsvektoren rekombiniert. Die verwendeten Primer 30-32 und 35-37 sind in Tabelle II.3 angegeben. Schrittweise konnte so *TDSCO-80* mit *gcn4tt-adh1p* verbunden und das Fusionsprodukt mit *GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc* verlinkt werden. Wie auch bei der Generierung des ersten Fusionskonstrukts wurde das erhaltene Fragment *TDSCO-80-gcn4tt-adh1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc* (*T-G-t GA*) via Gateway-Technologie in den Destinationsvektor pAG423 kloniert (pAG423::*T-G-t GA*). Die dafür benötigten Rekombinationssequenzen wurden auch hier bereits bei der Stargate-PCR integriert. Nach Bestätigung der korrekten Basenabfolge durch Sequenzierungen wurde das neue Fusionskonstrukt verwendet, um den CEN.PK2-1C-Stamm zu transformieren und der generierte Fusionsstamm CEN 10 mittels GC/MS-Messungen auf die Produktion von Taxadien, GGOH und Squalen hin untersucht. Dabei konnte kein Taxadien detektiert werden. Auch die Werte für GGOH und Squalen entsprachen trotz positiver Selektion des transformierten Stamms in Histidin-negativem Medium denen des Wildtypstamms. Eine Northernblotanalyse zeigte eine schwache Transkription von *TDSCO-80* (Abbildung III.20). Die Translation der gebildeten mRNS zu *TDSCO-80* konnte anhand eines Westernblots nicht nachgewiesen werden.



**Abbildung III.20: Northernblotanalyse des modifizierten Fusionsstamms. Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10 µg geladen. **Unten:** Autoradiogramm der spezifisch gebundenen radioaktiv markierten DNS-Sonde; **CEN 7:** Einzeln exprimierte Gene *TDSco*-, *GGPPS-Sa* und *tHMGRSc*; **CEN 10:** *TDSco-80*, *GGPPS-Sa* und *tHMGRSc* als Tandemkonstrukt angeordnet mit identischen internen Regulationseinheiten *gcn4tt-adh1p*; **Wt:** CENPK2-1C Wildtyp

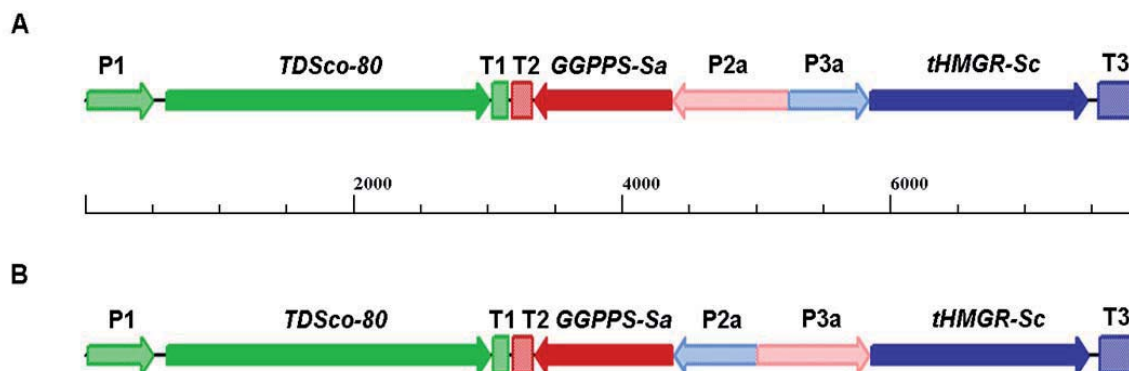
Dieser und alle anderen im Rahmen der Arbeit entstandenen Fusionsstämme sind in Tabelle III.8 zusammengefasst.

**Tabelle III.8: Im Rahmen dieser Arbeit generierte Fusionskonstrukte**

Stamm	Plasmid	Relevanter Genotyp
CEN 8	pAG423:: <i>T-G-t</i>	<i>TDSco-80-pho5TT-gap1P-GGPPS-Sa-gcn4TT-adh1P-tHMGR1</i>
CEN 9	pAG423:: <i>T-G-t inv</i>	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-gap1P-adh1P-tHMGR1</i>
CEN 10	pAG423:: <i>T-G-t GA</i>	<i>TDSco-80- gcn4TT-adh1P-GGPPS-Sa-gcn4TT-adh1P-tHMGR1</i>
CEN 11	pAG423:: <i>T-G-t P-T</i>	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-pgkP-tefP-tHMGR1</i>
CEN 12	pAG423:: <i>T-G-t T-P</i>	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-tefP-pgkP-tHMGR1</i>
CEN 13	pAG423:: <i>T-G-t ADH</i>	<i>TDSco-80-pho5TT-pgkP+-GGPPS-Sa-cyc1TT+-adh1P-tHMGR1</i>
CEN 14	pAG423:: <i>T-G-t TEF</i>	<i>TDSco-80-pho5TT-pgkP+-GGPPS-Sa-cyc1TT+-tefP-tHMGR1</i>
CEN 15	pAG423:: <i>T-L-G-t</i>	<i>TDSco-80-L- GGPPS-Sa-gcn4TT-adh1P-tHMGR1</i>
CEN 16	pAG423:: <i>T-G-t co</i>	<i>TDSco-80-pho5TT-tefP-GGPPSco-adh1TT-adh1P-tHMGR1</i>
CEN 17	pVV200:: <i>T-G-t co</i>	<i>TDSco-80-pho5TT-tefP-GGPPSco-adh1TT-adh1P-tHMGR1</i>
CEN 18	pVV200:: <i>T-G-t T-P</i>	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-tefP-pgkP-tHMGR1</i>

***TDSco-80*:** codon-optimierte Taxadiensynthase, *N*-terminal um 80 Aminosäuren verkürzt; ***GGPPS-Sa*:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; ***tHMGR1*:** Gensequenz für das HMG-CoA-Reduktase Isozyms 1 aus *Saccharomyces cerevisiae*, verkürzt um die *N*-terminale Regulationsdomäne; ***pho5TT*:** Terminator der sauren Phosphatase; ***gap1P*:** Allgemeine Aminosäurepermease-Promotor; ***gcn4TT*:** Terminator des Transkriptionsaktivators von Aminosäure-Biosynthesegenen; ***adh1P*:** Alkoholdehydrogenase-Promotor; ***tefP*:** Promotor des translationalen Elongationsfaktorgens; ***pgkP+*:** Phosphoglyceratkinase-Promotorregion aus pVV200 ; ***cyc1TT+*:** Cytochrom 1-Terminatorregion aus pVV200; ***L*:** Peptidlinker Gly-Gly-Gly-Ser; ***GGPPco*:** codon-optimiertes Geranylgeranyldiphosphatsynthasegen; ***adh1TT*:** Alkoholdehydrogenase-Terminator

Mehrere Hefeexpressionsvektoren wurden entwickelt, die erfolgreich den konstitutiven Promotoren des Translationselongationsfaktors (*tefp*) oder der Phosphoglyceratkinase (*pgkp*) verwenden<sup>[154][183][221]</sup> und bei ATCC, Euroscarf oder DualsystemsBiotech kommerziell erhältlich sind. Daher wurde ein weiteres Cluster konstruiert, in dem der bisher verwendete *gap1p* vor dem nicht transkribierten *GGPPS-Sa* ersetzt wurde. Das Fusionskonstrukt mit der konvergent-divergenten Orientierung zeigte die bisher beste Produktionsleistung (Tabelle III.7), sodass für die Modifizierung der Vektor *pAG423::T-G-t inv* (Abbildung III.16) verwendet wurde. Um die Promotorensequenz *gap1p-adh1p* gegen *pgkp-tefp* auszutauschen, wurden die *PGK*- und *TEF*-Promotoren über SOE-PCR fusioniert und phosphoryliert (Primer 40-43, Tabelle II.3). Das Fragment *GGPPS-Sa-gcn4tt-pho5tt-TDSco-80-pAG423-cyc1tt-tHMGR-Sc* wurde aus dem Vektor amplifiziert (Primer 38 und 39, Tabelle II.3) und mit dem SOE-PCR-Produkt *pgkp-tefp* ligiert. Die erhaltenen Fusionskonstrukte *pAG423::TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-pgkp-tefp-tHMGR-Sc* (*pAG423:: T-G-t P-T*) und *pAG423::TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-tefp-pgkp-tHMGR-Sc* (*pAG423::T-G-t T-P*) sind in Abbildung III.21 dargestellt.



**Abbildung III.21: Fusionskonstrukt *pAG423::T-G-t P-T* (A) und *pAG423::T-G-t T-P* (B).** Die Gene *TDSco-80* und *GGPPS-Sa* sind konvergent orientiert, *GGPPS-Sa* und *tHMGR-Sc* stehen divergent zueinander und werden über die Promotoreinheiten *pgkp-tefp* (A) bzw. *tefp-pgkp* (B) getrennt. **P1:** *GPD*-Promotor; **T1:** *PHO5*-Terminator; **P2:** *PGK*-Promotor; **T2:** *GCN4*-Terminator; **P3:** *TEF*-Promotor; **T3:** *CYC1*-Terminator

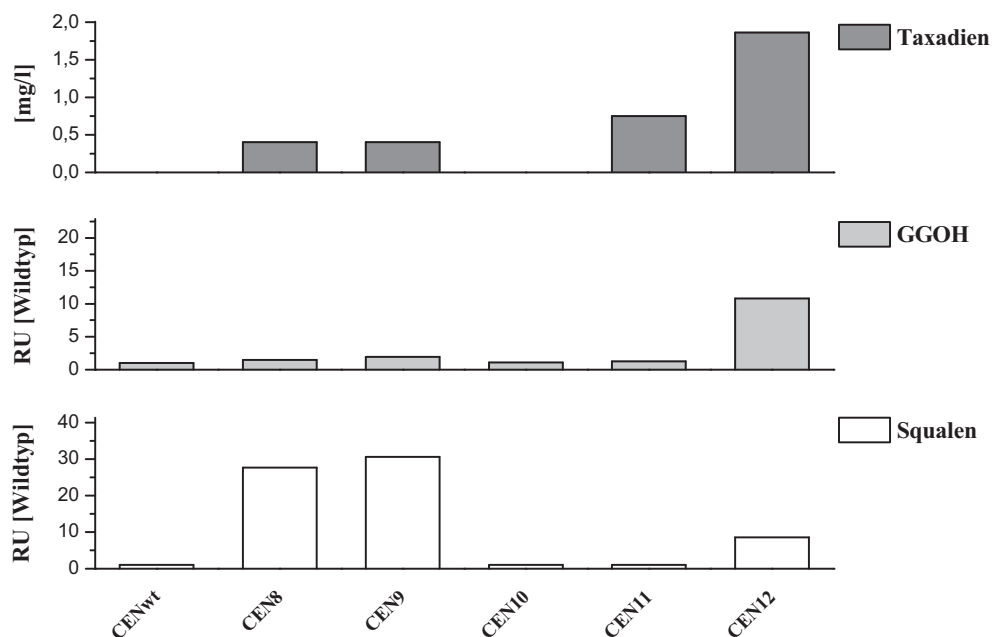
Sequenzierungen bestätigten die korrekte Basenabfolge beider Vektoren. Durch Transformation des *Saccharomyces cerevisiae*-Stamms CEN.PK2-1C mit den Vektoren wurden CEN 11 und CEN 12 generiert. Eine GC/MS-Analyse der Stämme zur Quantifizierung synthetisierten Taxadiens, GGOHs und Squalens erbrachte die in Tabelle III.9 angegebenen Produktmengen.

**Tabelle III.9: Produktanalyse modifizierter Fusionsstämme via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/L]	GGOH [mg/L]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>G</i> ; <i>t</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±5,6	3,90±0,3
CEN 8	<i>T-pho5tt-gap1p-G-gcn4tt-adh1p-t</i>	0,41±0,02	0,34±0,05	86,3±12
CEN 9	<i>Tpho5tt-gcn4tt-G-gap1p-adh1p-t</i>	0,42±0,005	0,45±0,03	94,8±21
CEN 10	<i>T-gcn4tt-adh1p-G-gcn4tt-adh1p-t</i>	n.d.	0,25±0,01	3,20±0,1
CEN 11	<i>T-pho5tt-gcn4tt-G-pgkp-tefp-t</i>	0,78±0,01	0,29±0,03	3,12±0,2
CEN 12	<i>T-pho5tt-gcn4tt-G-tefp-pgkp-t</i>	1,86±0,1	2,50±0,1	26,5±7

**T:** *TDSco-80*; **G:** *GGPPS-Sa*; **t:** *tHMGR-Sc*; **pho5TT:** Terminator der sauren Phosphatase **gap1P:** Allgemeine Aminosäurepermease-Promotor; **gcn4TT:** Terminator des Transkriptionsaktivators von Aminosäure-Biosynthesegenen; **adh1P:** Alkoholdehydrogenase-Promotor; **tefp:** Promotor des translationalen Elongationsfaktorgens

Unabhängig von der Orientierung der Regulationseinheiten *pgkp-tefp* innerhalb des Genclusters resultierte der Austausch der Promotoren verglichen mit den bisherigen Fusionsstämmen CEN 8-10 in den bisher höchsten Konzentrationen an Taxadien (Abbildung III.22).

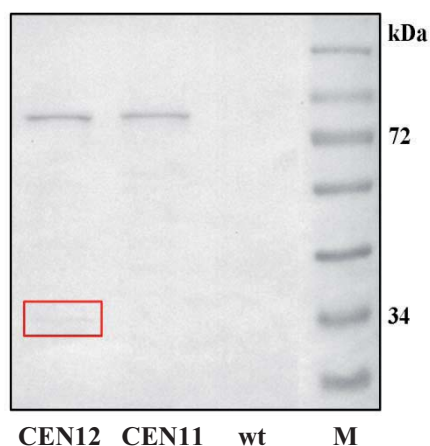


**Abbildung III.22: Vergleich generierter Fusionsstämme auf ihre Produktion an Taxadien, GGOH und Squalen mittels GC/MS-Analyse.** Taxdienkonzentration angegeben in mg/L, GGOH und Squalen jeweils normiert auf Wildtypwerte. RU: Relative Einheit bezogen auf den Wildtyp. **CEN 8:** transformiert mit pAG423::*TDSco-80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR1*; **CEN 9:** transformiert mit pAG423::*TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-gap1p-adh1p-tHMGR1*; **CEN 10:** transformiert mit pAG423::*TDSco-80-gcn4tt-adh1p-GGPPS-Sa-gap1p-adh1p-tHMGR1*; **CEN 11:** transformiert mit pAG423::*TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-pgkp-tefp-tHMGR1*; **CEN 12:** transformiert mit pAG423::*TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-tefp-pgkp-tHMGR1*



Wurde der *PGK*-Promotor vor *GGPPS-Sa* geschaltet, synthetisierte der entsprechende Hefestamm CEN 11 trotz leicht verminderter GGOH-Ausbeute (0,29 mg/L) mit 0,78 mg/L mehr als doppelt soviel Taxadien wie der initial generierte Fusionsstamm CEN 8. Bei Expression von *GGPPS-Sa* mithilfe des *TEF*-Promotors (CEN 12) erhöhte sich die Taxadienmenge nochmals auf über das Doppelte (1,86 mg/L). Die Bestimmung der GGOH-Konzentration von 2,50 mg/L ergab verglichen mit dem CEN 8 eine gleichzeitige Steigerung um das achtfache. Die Quantifizierung von Squalen demonstrierte, dass die gebildete Menge von dem vor *tHMGR1-Sc* lokalisierten Promotor abhängt. Der in CEN 11 dafür eingesetzte *tefp* führt zu einer mit dem Wildtyp vergleichbaren Squalenmenge von 3,12 mg/l. Wird hingegen der *pgkp* vor *tHMGR-Sc* ligiert (CEN 12), steigt die detektierbare Squalenkonzentration auf 26,5 mg/l.

Die unterschiedlichen Produktionsraten von CEN 11 und CEN 12 spiegeln sich auch in der Akkumulation der entsprechenden Proteine wider (Abbildung III.23). Während bei dem erfolgreicheren Produktionsstamm CEN 12 sowohl eine Bande für die TDSco-80, als auch eine für die GGPPS-Sa detektierbar war, fehlte letztere in der Proteinfraction von CEN 11. *tHMGR-Sc* dagegen war anhand eines Westernblots bei keinem der beiden Fusionsstäme nachweisbar.

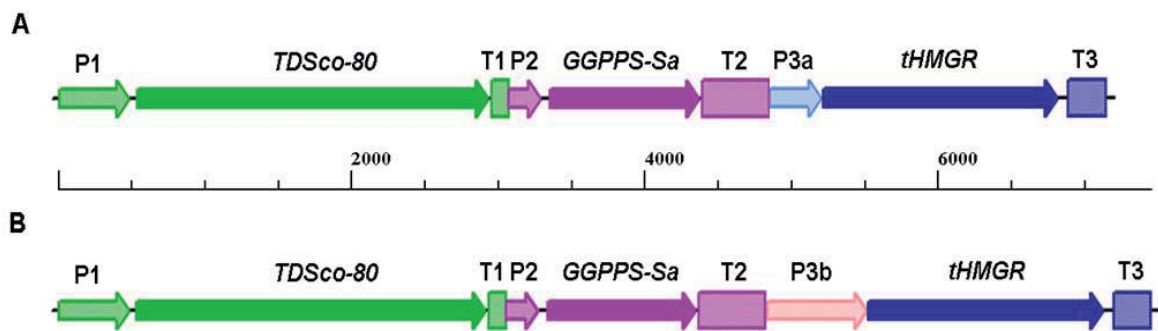


**Abbildung III.23: Westernblot von 40 µg TCA-gefällten Gesamtproteins generierter Fusionsstäme. CEN 12:** exprimiert Gencluster bestehend aus konvergent orientiertem *TDSco-80* und *GGPPS-Sa*, sowie divergent angeordneten *GGPPS-Sa* und *tHMGR-Sc*. Letztere getrennt durch *pgkp-tefp*; **CEN 11:** exprimiert Gencluster bestehend aus konvergent orientiertem *TDSco-80* und *GGPPS-Sa*, sowie divergent angeordneten *GGPPS-Sa* und *tHMGRSc*. Letztere getrennt durch *tefp-pgkp*; **Wt:** CENPK2-1C Wildtyp; **M:** Fermentas-Größenstandard

Trotz verbesserter Produktionsleistung des Fusionsstamms CEN 12 ist die synthetisierte Menge an Taxadien für die Umsetzung zu weiteren Intermediaten des Taxolbiosynthesewegs sehr begrenzt. Da *GGPPS-Sa* im Einzelvektor pVV200::*GGPPS-Sa* erfolgreich exprimiert werden konnte (Abbildung III.3), wurde als alternativer Optimierungsansatz des Fusionskonstrukts das Fragment *gap1p-GGPPS-Sa-gcn4tt* gegen das Gen inklusive *PGK*-Promotor und *CYC1*-Terminator aus dem Einzelvektor ausgetauscht. Das phosphorylierte PCR-Produkt (Primer 44 und 45, Tabelle II.3)



wurde anschließend mit dem aus pAG423::*T-G-t* **T-P** (Tabelle III.8) amplifizierten Vektorteil pho5tt-*TDSco-80*-pAG423-cyc1tt-*tHMGR-Sc-adh1p* (Primer 16 und 17, Tabelle II.3), bzw. aus pAG423::*T-G-t* **P-T** (Tabelle III.8) amplifizierten pho5tt-*TDSco-80*-pAG423-cyc1tt-*tHMGR-Sc-tefp* ligiert. Die Inserts der neuen Fusionskonstrukte pAG423::*T-G-t* ADH und pAG423::*T-G-t* TEF (Tabelle III.8) sind in Abbildung III.24 abgebildet.



**Abbildung III.24: Schematische Darstellung der Inserts von Fusionskonstrukt pAG423::*T-G-t* ADH (A) und pAG423::*T-G-t* TEF (B).** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind im Tandem orientiert. Die Einheit pgkp-*GGPPS-Sa*-cyc1tt wurde aus dem Einzelvektor pVV200::*GGPPS-Sa* amplifiziert und anstelle von gcn4tt-*GGPPS-Sa*-tefp in pAG423::*T-G-t* T-P (A), bzw. anstelle von gcn4tt-*GGPPS-Sa*-pgkp in pAG423::*T-G-t* P-T (B) eingesetzt. ***TDSco-80***: codon-optimiertes Gen für Taxadiensynthese, N-terminal um 80 AS verkürzt; **T1-3**: Terminatorsequenzen der gleichfarbig dargestellten Gene; **P1-3**: Promotorsequenzen der gleichfarbig dargestellten Gene

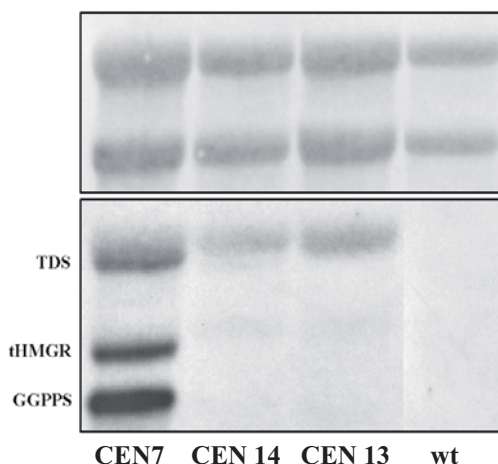
Die korrekte Basenabfolge der Fusionsvektoren pAG423::*T-G-t* ADH und pAG423::*T-G-t* TEF (Tabelle III.8) wurde mittels Sequenzierungen bestätigt und durch die Transformation des *S. cerevisiae*-Stamms CEN.PK2-1C die Stämme CEN 13 und CEN 14 generiert. Anhand von GC/MS-Messungen wurden diese auf produzierte Mengen an Taxadien, GGOH und Squalen hin untersucht. In Tabelle III.10 sind die jeweiligen Quantifizierungsergebnisse im Vergleich mit den bisher untersuchten Fusionsstämmen angegeben.

**Tabelle III.10: Produktanalyse generierter Fusionsstämme via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>G</i> ; <i>t</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±5,6	3,90±0,3
CEN 8	<i>T-pho5tt-gap1p-G-gcn4tt-adh1p-t</i>	0,41±0,02	0,34±0,05	86,3±12
CEN 9	<i>Tpho5tt-gcn4tt-G-gap1p-adh1p-t</i>	0,42±0,005	0,45±0,03	94,8±21
CEN 10	<i>T-gcn4tt-adh1p-G-gcn4tt-adh1p-t</i>	n.d.	0,25±0,01	3,20±0,1
CEN 11	<i>T-pho5tt-gcn4tt-G-pgkp-tefp-t</i>	0,78±0,01	0,29±0,03	3,12±0,2
CEN 12	<i>T-pho5tt-gcn4tt-G-tefp-pgkp-t</i>	1,86±0,1	2,50±0,1	26,5±7
CEN 13	<i>T-pho5tt-pgkp+-G-cyc1tt+-adh1p-t</i>	0,41±0,008	4,9±0,1	47,1±2,7
CEN 14	<i>T-pho5tt-pgkp+-G-cyc1tt+-tefp-t</i>	n.d.	0,23±0,007	3,30±0,1

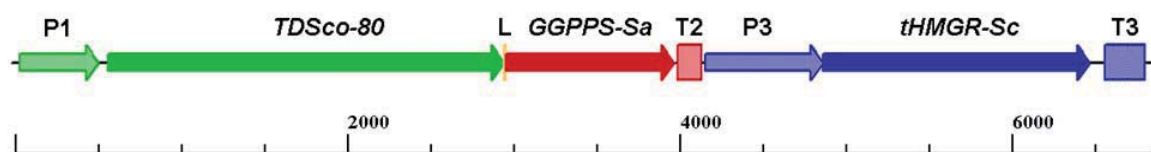
**T:** *TDSco-80*; **G:** *GGPPS-Sa*; **t:** *tHMGR-Sc*; **pho5TT:** Terminator der sauren Phosphatase **gap1P:** Allgemeine Aminosäurepermease-Promotor; **gcn4TT:** Terminator des Transkriptionsaktivators von Aminosäure-Biosynthesegenen; **adh1P:** Alkoholdehydrogenase-Promotor; **tefp:** Promotor des translationalen Elongationsfaktorgens; **pgkP+:** Phosphoglyceratkinase-Promotorregion aus pVV200; **cyc1TT+:** Cytochrom c-Terminatorregion aus pVV200

Die Quantifizierung des von CEN 13 gebildeten Taxadiens ergab eine Ausbeute von 0,41 mg/L, was denen der ersten Fusionsstämme CEN 8 und 9 entspricht (Tabelle III.10). Die Produktion von GGOH hingegen lag mit 4,9 mg/L bei dem ca. zehnfachen Wert, die Squalenmenge mit 47,1 mg/l ungefähr bei der Hälfte. Trotz durch Sequenzierung bestätigtem, korrektem Insert des Vektors pAG423::*T-G-t* TEF (Tabelle III.8) und Selektion des damit generierten CEN 14 auf die entsprechende Aminosäurekomplementation, entsprachen die bestimmten Produktionswerte für Taxadien, GGOH und Squalen den jeweiligen des Wildtypstamms. Ein Nachweis synthetisierter mRNS der einzelnen Clustergene konnte bei beiden Stämmen ausschließlich für *TDSco-80* erbracht werden (Abbildung III.25).



**Abbildung III.25: Northernblotanalyse generierter Fusionsstämme.** **Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10 µg geladen. **Unten:** Autoradiogramm der spezifisch gebundenen, radioaktiv markierten DNS-Sonde; **CEN7:** Einzeln exprimierte Gene *TDSco-*, *GGPPS-Sa* und *tHMGRSc*; **CEN 13 und 14:** *TDSco-80*, *GGPPS-Sa* und *tHMGRSc* als Tandemkonstrukt angeordnet. *GGPPS-Sa* inklusive Promotor und Terminator wurden aus dem Einzelvektor pVV200::*GGPPS-Sa* amplifiziert. Promotor vor *tHMGR-Sc* *adh1p* (CEN 13) oder *tefp* (CEN 14); **Wt:** CENPK2-1C Wildtyp.

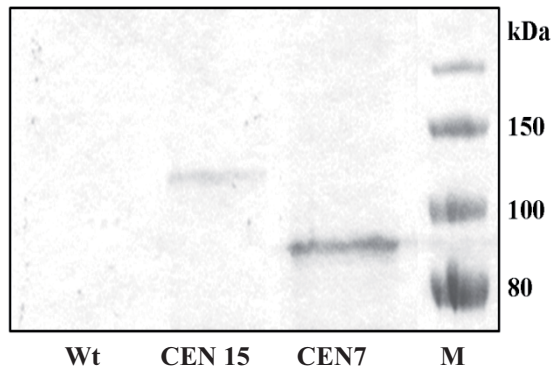
Die sieben *S. cerevisiae*-Fusionsstämme CEN 8-14 sind durch die Transformation mit verschiedenen Fusionsvektoren generiert worden (Tabelle III.8). Diese enthielten jeweils die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc*, die in unterschiedlicher Orientierung und mit wechselnden Promotoren, bzw. Terminatoren verbunden waren. Auch wenn bei manchen der generierten Stämme eine geringe Taxadienproduktion zu beobachten war, entsprach diese doch nicht derjenigen des die Einzelgene exprimierenden Stamms CEN 7. Anhand von Northern- und Westernblots konnte das Problem auf die fehlende Expression des mittig lokalisierten Gens *GGPPS-Sa* zurückgeführt werden. Um die Problematik der Regulationseinheiten von *GGPPS-Sa* zu umgehen, wurde ein Cluster generiert, in dem *TDSco-80* und *GGPPS-Sa* von dem gleichen *GPD*-Promotor reguliert werden und für ein Fusionsprotein kodieren (pAG423::*T-G-t* L, Abbildung III.26). *TDSco-80* und *GGPPS-Sa* sind dabei über den Peptidlinker Gly-Gly-Gly-Ser verbunden. Dieser wurde bereits erfolgreich zur Fusionierung von GGPPS mit der Diacylglyceroldiphosphat-Phosphatase DPP1 eingesetzt.<sup>[222]</sup>



**Abbildung III.26: Schematische Darstellung der Inserts von Fusionskonstrukt pAG423::*T-G-t* L.** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind im Tandem orientiert. *TDSco-80* und *GGPPS-Sa* sind über Gly-Gly-Gly-Ser (L) miteinander verbunden und kodieren für ein Fusionsprotein. **P1:** *GPD*-Promotor; ***TDSco-80*:** codon-optimiertes Gen für Taxadien-synthese, N-terminal um 80 AS verkürzt; ***GGPPS-Sa*:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **T2:** *GCN4*-Terminator; **P3:** *ADH1*-Promotor; ***tHMGR-Sc*:** Gen kodierend für die katalytische Domäne des HMGcoA-Reduktase Isozyms 1 aus *S. cerevisiae*; **T3:** *CYC1*-Terminator

Die Verlinkung von *TDSco-80* und *GGPPS-Sa* erfolgte mittels SOE-PCR. Die beiden Fragmente *TDSco-80* und *GGPPS-Sa-gcn4tt-adh1p-tHMGR-Sc* wurden dafür aus dem initialen Vektor pAG423::*T-G-t* amplifiziert (Primer 6 und 44-46, Tabelle II.3). Das Produkt wurde anschließend mithilfe der Gateway-Technologie über die endständig angefügten Erkennungsstellen in den Expressionsvektor pAG423 rekombiniert. Mit dem generierten Vektor pAG423::*T-G-t* L wurde der Hefestamm CEN.PK2-1C transformiert und der entstandene Stamm CEN 15 (Tabelle III.8) auf eine aktive Expression der Gene untersucht. Obgleich ein Nachweis des Fusionsproteins anhand eines

Westernblots erbracht werden konnte (Abbildung III.27), ergab die Produktanalyse des Stamms mittels GC/MS keine detektierbaren Mengen an gebildetem Taxadien.



**Abbildung III.27: Ausschnitt aus Westernblot von 40 µg TCA-gefällten Gesamtproteins des generierten Linker-Fusionsstamms CEN 15. Wt: CENPK2-1C Wildtyp; CEN 15: exprimiert Gencluster bestehend aus *TDSco-80* und *GGPPS-Sa* als Fusionsprotein, sowie im Tandem dazu angeordnetem *tHMGRSc*. CEN 7: exprimiert *TDSco*, *GGPPS-sa* und *tHMR-Sc* in Einzelvektoren. M: Fermentas-Größenstandard**

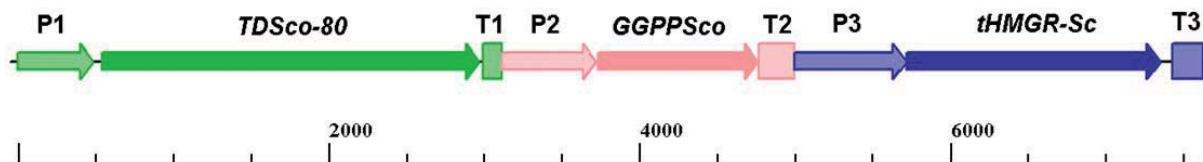
Bei einer GGOH-Konzentration von 0,64 mg/L konnte allerdings mit 119,4 mg/l eine dem Wildtyp gegenüber auf das 40fache gesteigerte Menge an Squalen nachgewiesen werden (Tabelle III.11).

**Tabelle III.11: Quantifizierung ausgewählter Produkte generierter Fusionsstämme via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±5,6	3,90±0,3
CEN 8	<i>T-pho5tt-gap1p-G-gcn4tt-adh1p-t</i>	0,41±0,02	0,34±0,05	86,3±12
CEN 9	<i>Tpho5tt-gcn4tt-G-gap1p-adh1p-t</i>	0,42±0,005	0,45±0,03	94,8±21
CEN 10	<i>T-gcn4tt-adh1p-G-gcn4tt-adh1p-t</i>	n.d.	0,25±0,01	3,20±0,1
CEN 11	<i>T-pho5tt-gcn4tt-G-pgkp-tefp-t</i>	0,78±0,01	0,29±0,03	3,12±0,2
CEN 12	<i>T-pho5tt-gcn4tt-G-tefp-pgkp-t</i>	1,86±0,1	2,50±0,1	26,5±7
CEN 13	<i>T-pho5tt-pgkp+-G-cyc1tt+-adh1p-t</i>	0,41±0,008	4,90±0,1	47,1±2,7
CEN 14	<i>T-pho5tt-pgkp+-G-cyc1tt+-tefp-t</i>	n.d.	0,23±0,007	3,30±0,1
CEN 15	<i>T-L-G-gcn4tt-adh1p-t</i>	n.d.	0,64±0,009	119±24
CEN 16	<i>T-pho5tt-tefp-Gco-adh1tt-adh1p-t</i>	0,46±0,005	0,36±0,007	4,00±0,1

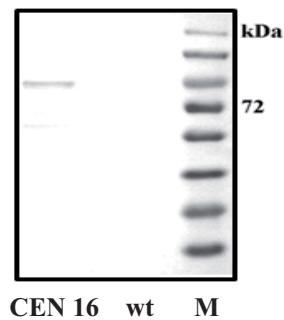
**T:** *TDSco-80*; **G:** *GGPPS-Sa*; **t:** *tHMGR-Sc*; **pho5TT:** Terminator der reprimierbaren Säurephosphatase **gap1P:** Genereller Aminosäure Permease-Promotor; **gcn4TT:** Terminator des Transkriptionsaktivators von Aminosäure-Biosynthesegenen; **adh1P:** Alkoholdehydrogenase-Promotor; **tefp:** Promotor des translationalen Elongationsfaktorgens; **pgkP+:** Phosphoglyceratkinase-Promotorregion aus pVV200; **cyc1TT+:** Cytochrom c-Terminatorregion aus pVV200; **L:** Peptidlinker Gly-Gly-Gly-Ser; **Gco:** für *S. cerevisiae* codon-optimiertes Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **adh1tt:** Terminatorregion des Alkoholdehydrogenasegens

Obwohl das bisher verwendete archaebakterielle *GGPPS-Sa* in *Saccharomyces cerevisiae* als Einzelgen aktiv exprimiert und translatiert werden konnte (Abbildung III.3), war es doch in allen Konstrukten diejenige Gensequenz, die zu keiner mittels Northernblot nachweisbaren Expression führte. Um eine mögliche Beeinträchtigung des heterologen Gens innerhalb eines Clusters aufgrund seiner ursprünglichen Sequenz auszuschließen, wurde eine abweichende, für *S. cerevisiae* codon-optimierte Version (*GGPPSco*) inklusive *TEF*-Promotor und *ADH1*-Terminator synthetisiert (Genescript) und zwischen *TDSco-80-pho5tt* und *adh1p-tHMGR-Sc* des initialen Fusionsvektors pAG423::*T-G-t* positioniert (Abbildung III.28).



**Abbildung III.28: Schematische Darstellung des Inserts von Fusionskonstrukt pAG423::*T-G-t* co.** Die Gene *TDSco-80*, *GGPPSco* und *tHMGR-Sc* sind im Tandem orientiert. *GGPPSco* inklusive Promotor und Terminator wurden synthetisch hergestellt. **P1:** *GPD*-Promotor; **T1:** *PHO5*-Terminator; **P2:** *TEF*-Promotor; **T2:** *ADH1*-Terminator; **P3:** *ADH1*-Promotor; **T3:** *CYC1*-Terminator

Die Integration erfolgte anhand einer Ligation des phosphorylierten Inserts *tefp-GGPPSco-adh1tt* mit dem amplifizierte Vektor *pho5tt-TDSco-80-pAG423-tHMGR-Sc-adh1p*. Die verwendeten Primer 16 und 17 sind in Tabelle II.3 angegeben. Mit dem entstandenen Fusionsvektor pAG423::*T-G-t* co (Tabelle III.8) wurde CEN.PK2-1C transformiert und der generierte Stamm CEN 16 via GC/MS auf seine Produktionsleistung an Taxadien, GGOH und Squalen untersucht. Die erhaltenen Werte sind in Tabelle III.11 abzulesen. Die Expression des codon-optimierten GGPPS-Gens (CEN 16) führte verglichen mit dem ersten Fusionsstamm CEN 8 zu leicht erhöhten Werten gebildeten Taxadiens und GGOHs ( $456 \pm 5$ , bzw.  $356 \pm 7$   $\mu\text{g/l}$  gegenüber  $405 \pm 17$ , bzw.  $338 \pm 47$   $\mu\text{g/l}$ ). Die Konzentration an detektierbarem Squalen dagegen zeigte sich mit  $4 \pm 0,1$  mg/l ähnlich zu der in CEN 7 akkumulierten Menge. Zur weiteren Analyse wurde aus CEN 16 Gesamtprotein extrahiert und im Vergleich mit der entsprechenden Fraktion aus dem Wildtypstamm CEN.PK2-1C via Westernblot analysiert (Abbildung III.29).



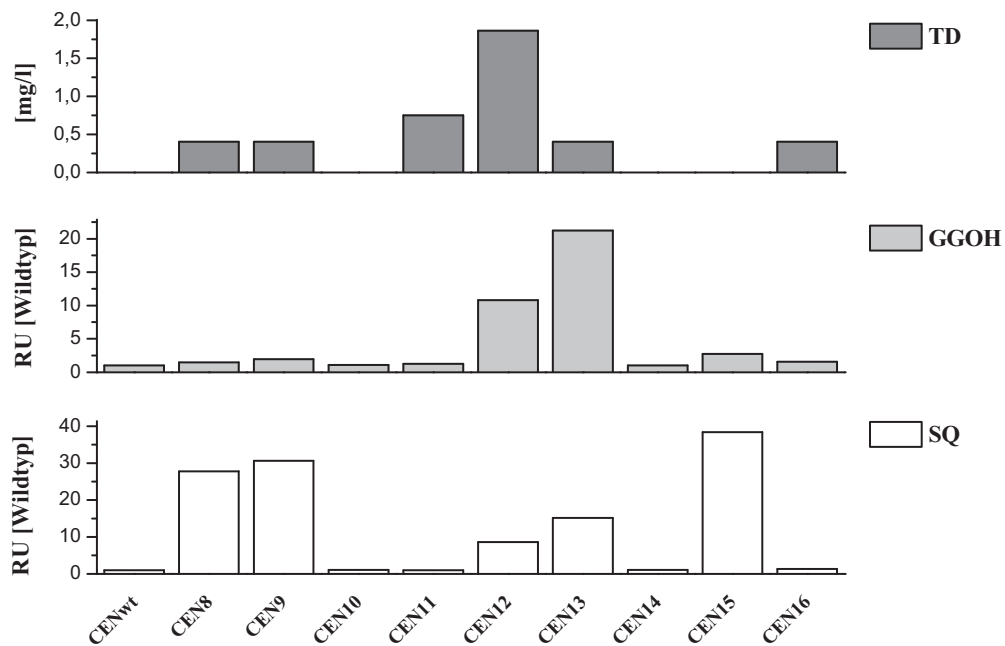
**Abbildung III.29:** Westernblot von 40 µg TCA-gefällten Gesamtproteins des generierten Fusionsstamms CEN 16. **Wt:** CENPK2-1C Wildtyp; **CEN 16:** exprimiert Gencluster bestehend aus *TDSco-80*, *GGPPSco* und *tHMGR-Sc* im Tandem angeordnet. **M:** Fermentas Größen-standard

Obschon alle Gene des Clusters Pag423::*T-G-t* co anhand von Sequenzierungen auf eine Histidinmarkierung überprüft worden waren, konnte bei Analyse der Proteinfraction von CEN 16 ausschließlich eine Bande in Höhe der Taxadiensynthese detektiert werden.

### III.3.4 Stabilitätsanalyse transformierter Fusionskonstrukte

Neun Konstrukte sind generiert worden, in denen die drei Gene *TDSco-80*, *GGPPS-Sa* (bzw. *GGPPSco*) und *tHMGR-Sc* fusioniert worden waren (Tabelle III.8). Trotz teilweiser Steigerung gegenüber dem Wildtyp (Abbildung III.30), führte nach Transformation von CEN.PK2-1C keines dieser Konstrukte zu der Produktionsleistung an Taxadien oder GGOH, die durch die Expression der Einzelgene beobachtet worden war. Weder die Einführung unterschiedlicher Promotoren und Terminatoren in das Cluster, noch eine variierende Orientierung der Gene zueinander resultierte in einer mittels Northernblot nachweisbaren Transkription aller drei Gene. Um zu überprüfen, ob die Stabilität der Fusionsplasmide in der Hefezelle das Problem darstellt und nicht das Konstrukt selbst, wurden die generierten Fusionsstämme sowohl auf den Verlust der Plasmide, als auch auf eine Veränderung derselben analysiert.





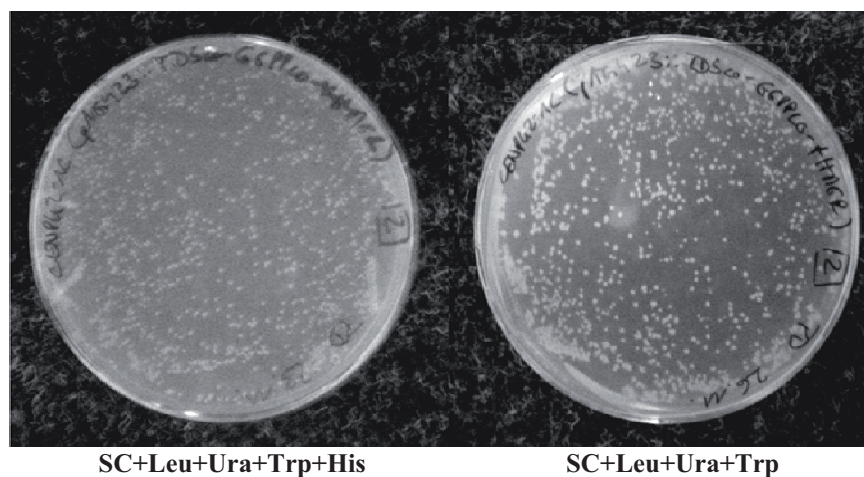
**Abbildung III.30: Vergleich der Produktion an Taxadien, GGOH und Squalen generierter Fusionsstämme mittels GC/MS-Analyse.** TD: Taxadien; GGOH: Geranylgeraniol; SQ: Squalen. Taxadienkonzentration angegeben in mg/l, GGOH und Squalen jeweils normiert auf Wildtypwerte. RU: Relative Einheit bezogen auf den Wildtyp. **CEN 8:** transformiert mit pAG423::*TDSCO-80-pho5tt-gap1p-GGPPS-Sa-gcn4tt-adh1p-tHMGR1*; **CEN 9:** transformiert mit pAG423::*TDSCO-80-pho5tt-gcn4tt-GGPPS-Sa-gap1p-adh1p-tHMGR1*; **CEN 10:** transformiert mit pAG423::*TDSCO-80-gcn4tt-adh1p- GGPPS-Sa-gap1p-adh1p-tHMGR1*; **CEN 11:** transformiert mit pAG423::*TDSCO-80-pho5tt- gcn4tt-GGPPS-Sa-pgkp-tefp-tHMGR1*; **CEN 12:** transformiert mit pAG423::*TDSCO-80-pho5tt- gcn4tt-GGPPS-Sa-tefp-pgkp-tHMGR1*; **CEN 13:** transformiert mit pAG423::*TDSCO-80-pho5tt-pgkp+-GGPPS-Sa-cyc1tt+-adh1p-tHMGR1*; **CEN 14:** transformiert mit pAG423::*TDSCO-80-pho5tt-pgkp+-GGPPS-Sa-cyc1tt+-tefp-tHMGR1*; **CEN 15:** transformiert mit pAG423::*TDSCO-80-L-GGPPS-Sa-gcn4tt-adh1p-tHMGR1*; **CEN 16:** transformiert mit pAG423::*TDSCO-80-pho5tt-tefp-GGPPS-Sa-adh1tt-adh1p-tHMGR1*

### III.3.4.1 Untersuchung auf Plasmidverlust

Die durchgeführte Produktanalyse aller generierten Fusionsstämme mittels GC/MS basiert auf einer Cyclohexanextraktion der jeweiligen Kultur nach 72 h Inkubation (II.1.3.2). Dieser Zeitraum unterteilt sich in 24 h selektiven Wachstums im entsprechenden SC-Medium (II.1.2.4) gefolgt von 48 h Kultivierung in Komplexmedium (II.1.2.3), in denen die Zellen zwar schneller wachsen, aber keinerlei Selektionsdruck unterliegen. Diese Kombination von selektiven und nicht-selektiven Anzuchtbedingungen war im Rahmen der Generierung des Taxadien-produzierenden Stamms CEN 7 (III.2) erfolgreich angewendet worden. Ein Verlust der Plasmide durch die Inkubation in Komplexmedium konnte weiterhin anhand von Stabilitätsuntersuchungen der entsprechenden Leervektoren für den gewählten Zeitraum ausgeschlossen werden. Die Daten waren im Rahmen



einer Diplomarbeit erstellt worden und zeigten eine Vektorstabilität zwischen 80 und 100% über einen Zeitraum von 50 h unter nicht-selektiven Bedingungen. Für die Generierung des Taxadien produzierenden Hefestamms CEN 7 waren mit pVV214, pVV200, pRS313 und pRS315 allerdings von dem für die Gencluster verwendeten Vektorgrundgerüst (pAG423) abweichende Leervektoren verwendet worden. Um die Möglichkeit auszuschließen, dass die Wahl des eingesetzten Vektors aufgrund seines Verlusts in der Hefezelle für die niedrige Produktionsleistung der Fusionsstämme verantwortlich ist, wurde pAG423 ebenfalls einer Stabilitätsuntersuchung unterzogen. Dazu wurde CEN 16 stellvertretend für alle generierten Fusionsstämme auf eine SC-Agarplatte mit Leucin, Uracil, Tryptophan und Histidin ausplattiert und die Kolonien anhand eines Stempels nach 48 h auf eine selektive SC-Agarplatte mit Leucin, Uracil und Tryptophan überführt. Nach 48 h Inkubationszeit wurde die Anzahl der Kolonien auf der Selektionsplatte mit derjenigen auf der nicht-selektiven Agarplatte verglichen (Abbildung III.31).

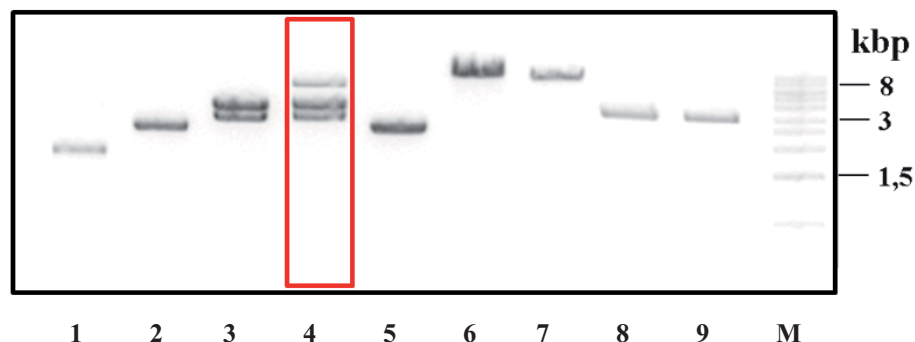


**Abbildung III.31: Stabilitätsanalyse des Fusionsvektors pAG423::T-G-t co.** Links: Wachstum des Fusionsstamms CEN 16 auf nicht-selektiver Agarplatte. Rechts: Wachstum von auf Selektionsplatte überführten Kolonien

Der Vergleich des Wachstums von CEN 16 unter nicht-selektiven Bedingungen (Abbildung III.31 links) mit demjenigen unter Selektionsbedingung (Abbildung III.31 rechts) zeigt nach 48 h eine ähnliche Anzahl an Kolonien. Ein Verlust des Fusionsplasmids konnte daher als Grund für fehlende oder niedrige Produktionsleistung der Fusionsstämme ausgeschlossen werden.

### III.3.4.2 Untersuchung auf Plasmidveränderung

Da ein fehlender Expressionsnachweis der fusionierten Gene nicht durch den Verlust der Konstrukte zu erklären ist (III.3.4.1), könnte eine Veränderung der Fusionsplasmide selbst der Grund für nicht exprimierte Gene innerhalb des Clusters sein. *Saccharomyces cerevisiae* ist bekannt für seinen hoch effizienten homologen Rekombinationsmechanismus, weshalb bereits diverse Genklonierungen,<sup>[185]</sup> Plasmidkonstruktionen<sup>[223]</sup> und Erstellungen von Bibliotheken<sup>[224]</sup> in diesem Organismus durchgeführt, sowie ganze Genome über homologe *in vivo* Rekombination in Hefe zusammengesetzt wurden.<sup>[225][226]</sup> Um Veränderungen von im Rahmen dieser Arbeit generierten Fusionskonstrukten innerhalb der Zelle detektieren zu können, wurden diese aus den Fusionsstämmen reextrahiert. Da die geringe Konzentration eine Analyse unmöglich machte, wurden die Plasmidextrakte aus der Hefe zur Amplifikation in *E. coli*-Zellen transformiert. Aus jeweils einer Kolonie wurde anschließend Plasmid-DNS isoliert und einer Restriktion mit der Endonuklease *Xba*I unterzogen. Die gelelektrophoretische Auftrennung der Ansätze ist in Abbildung III.32 dargestellt.



**Abbildung III.32: Gelelektrophoretische Auftrennung der mit *Xba*I verdauten, reextrahierten Fusionsplasmide.** Extrahiertes Plasmid aus Fusionsstamm **1:** CEN 8 (erwartetes Restriktionsmuster: 6,2 kb + 3,8 kb + 3,1 kb); **2:** CEN 9 (erwartetes Restriktionsmuster: 6,2 kb + 3,8 kb + 3,1 kb); **3:** CEN 10 (erwartetes Restriktionsmuster: 6,5 kb + 3,8 kb + 3,1 kb); **4:** CEN 11 (erwartetes Restriktionsmuster: 6,6 kb + 3,8 kb + 3,1 kb); **5:** CEN 12 (erwartetes Restriktionsmuster: 6,6 kb + 3,8 kb + 3,1 kb); **6:** CEN 13 (erwartetes Restriktionsmuster: 3,8 kb + 3,1 kb + 2,8 kb + 2,4 kb + 1,1 kb); **7:** CEN 14 (erwartetes Restriktionsmuster: 3,8 kb + 3,1 kb + 2,8 kb + 2,3 kb + 1,1 kb); **8:** CEN 15 (erwartetes Restriktionsmuster: 5,6 kb + 3,8 kb + 3,1 kb); **9:** CEN 16 (erwartetes Restriktionsmuster: 3,8 kb + 3,7 kb + 3,1 kb + 2,7 kb) **M:** Größenstandard NEB

Das Enzym *Xba*I wurde für den Testverdau der verschiedenen Fusionsplasmide gewählt, da die entsprechenden Erkennungssequenzen an Stellen innerhalb des Vektorgrundgerüsts (2micron, bzw. *GPD*-Promotor), sowie der Gensequenz von *tHMGR-Sc* schneidet, die in allen Konstrukten vorhanden sind. Jeder Restriktionsansatz sollte daher trotz variierender Promotor-/Terminatorsequenzen und unabhängig von der Orientierung der Gene zueinander die Fragmente

der Größe 3,1kb und 3,9kb für das geteilte Plasmidgrundgerüst enthalten. Bei Verdau der reextrahierten Fusionskonstrukte aus CEN 8-12 und CEN 16 wurde weiterhin ein zusätzliches Fragment von ungefähr 6 kb erwartet, das durch die unterschiedlichen Promotorlängen der einzelnen Konstrukte in seiner Größe leicht variieren kann. Dieses wird bei den Konstrukten aus CEN 13 und 14 zusätzlich geschnitten. Die gelelektrophoretische Auftrennung der Testrestriktionen ergab allerdings ein sehr unterschiedliches Fragmentierungsmuster der jeweiligen Ansätze (Abbildung III.32). Ausschließlich der *XbaI*-Verdau vom reextrahierten pAG423::*T-G-t* P-T aus CEN 11 zeigte die korrekten Bandengrößen von 6,4kb; 3,9kb und 3,1kb. Das Konstrukt pAG423::*T-G-t* GA aus CEN 10 teilte sich ausschließlich in die zwei vektorbasierenden Fragmente und scheint über kein Insert mehr zu verfügen. Alle anderen Konstrukte wurden durch *XbaI* linearisiert, wobei pAG423::*T-G-t* ADH aus CEN 13 und pAG423::*T-G-t* TEF aus CEN 14 zumindest die erwartete Gesamtgröße zeigten. Um die Veränderung der Plasmide innerhalb der Hefezellen genauer zu analysieren, wurden die in Tabelle III.12 aufgelisteten Konstrukte erneut in den *S. cerevisiae*-Stamm CEN.PK2-1C transformiert. Zu verschiedenen Zeitpunkten (nach 24, 48 und 72 h) wurde aus der Hefekultur Plasmid-DNS isoliert und diese mittels Transformation von *E. coli* vereinzelt. Aus jeweils drei Kolonien jedes Zeitpunktes wurde schließlich Vektor-DNS isoliert und mit *XbaI* verdaut.

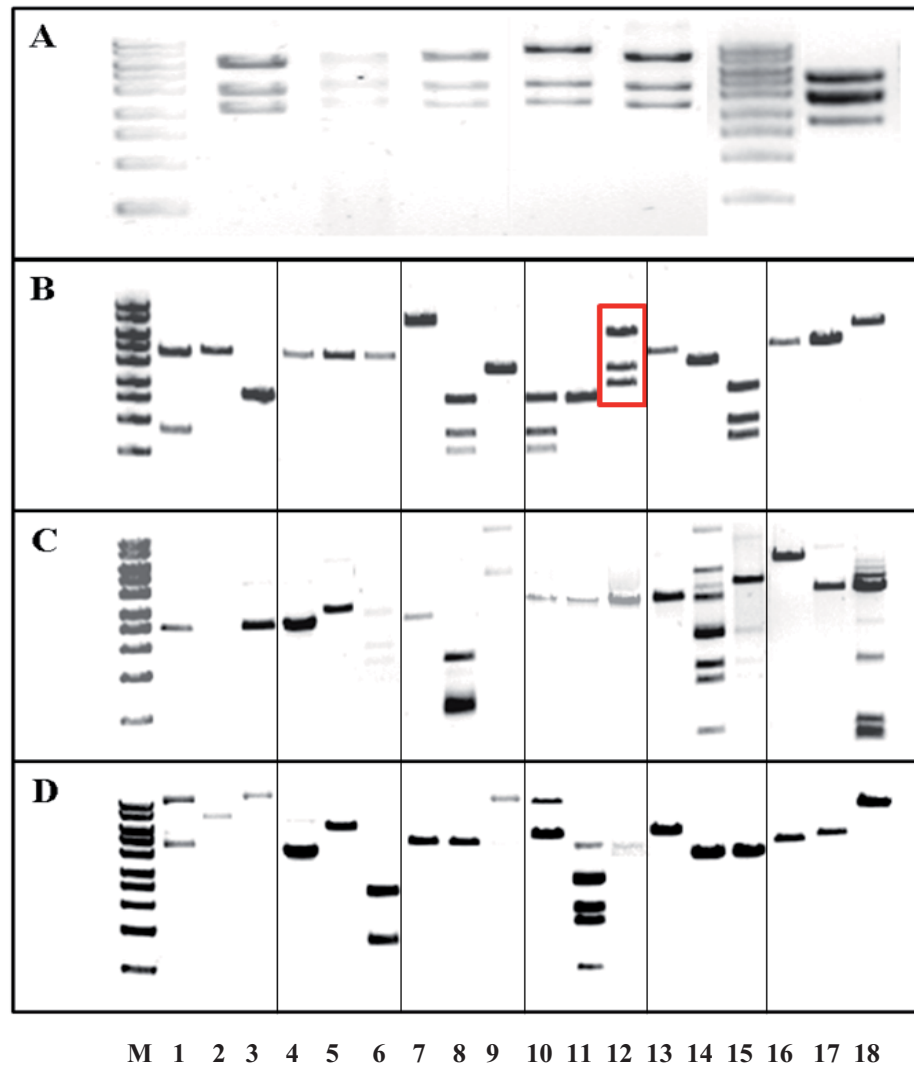
**Tabelle III.12: Auf Veränderungen untersuchte Fusionskonstrukte**

Plasmid	Relevanter Genotyp
pAG423:: <i>T-G-t</i>	<i>TDSco-80-pho5TT-gap1P-GGPPS-Sa-gcn4TT-adh1P-tHMGR1</i>
pAG423:: <i>T-G-t</i> inv	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-gap1P-adh1P-tHMGR1</i>
pAG423:: <i>T-G-t</i> GA	<i>TDSco-80- gcn4TT-adh1P-GGPPS-Sa-gcn4TT-adh1P-tHMGR1</i>
pAG423:: <i>T-G-t</i> P-T	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-pgkP-tefP-tHMGR1</i>
pAG423:: <i>T-G-t</i> T-P	<i>TDSco-80-pho5TT-gcn4TT-GGPPS-Sa-tefP-pgkP-tHMGR1</i>
pAG423:: <i>T-G-t</i> co	<i>TDSco-80-pho5TT-tefP-GGPPSco-adh1TT-adh1P-tHMGR1</i>

***TDSco-80***: codon-optimiertes Taxadiensynthasegen, kodierend für das *N*-terminal um 80 Aminosäuren verkürzte Protein; ***GGPPS-Sa***: Geranylgeranyldiphosphatsynthasegen aus *S. acidocaldarius*; ***tHMGR1***: Gensequenz der katalytischen Domäne des HMG-CoA-Reduktase Isozyms 1 aus *S. cerevisiae*; ***pho5TT***: Terminator der sauren Phosphatase; ***gap1P***: Allgemeine Aminosäurepermease-Promotor; ***gcn4TT***: Terminator des Transkriptions-aktivators von Aminosäure-Biosyntheseegenen; ***adh1P***: Alkoholdehydrogenase-Promotor; ***tefP***: Promotor des translationalen Elongationsfaktorgens; ***GGPPco***: codon-optimiertes Geranylgeranyldiphosphatsynthasegen; ***adh1TT***: Alkoholdehydrogenase-Terminator

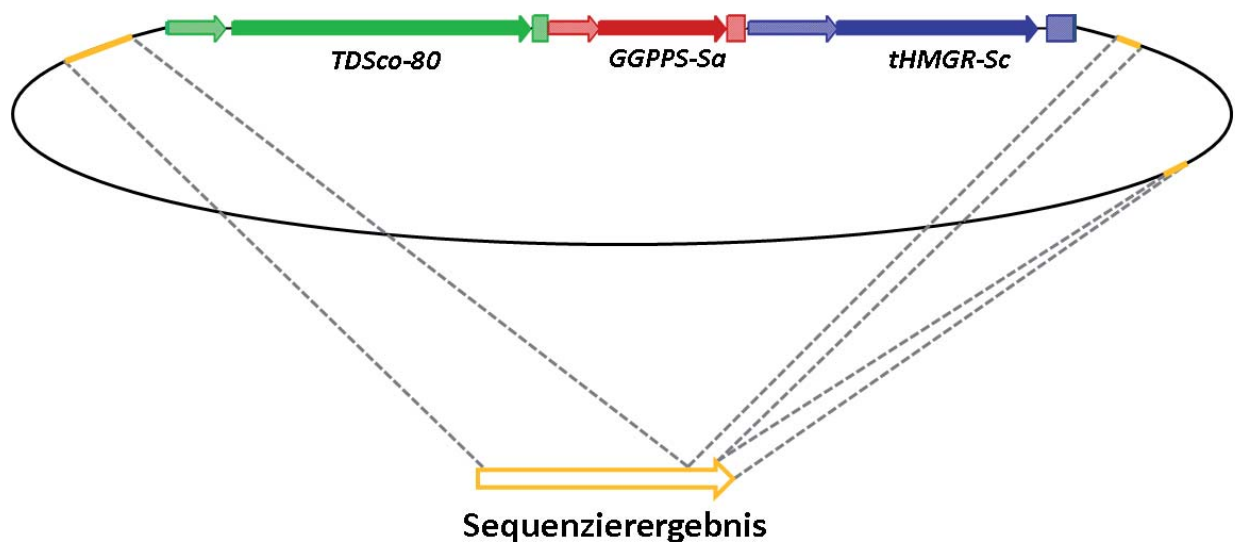
Die Zeitreihe sollte darüber Aufschluss geben, wann die Veränderung auftritt und ob sie ein kontinuierlicher Prozess ist, durch den sich die Fusionskonstrukte stetig ändern. Eine

Dreifachbestimmung sollte des Weiteren klären, ob die Veränderungen einheitlich für die einzelnen Konstrukte auftreten. Abbildung III.33 zeigt die gelelektrophoretische Auftrennung der *Xba*I-verdauten Plasmide vor der Transformation in die Hefezellen (A), sowie die reextrahierten Plasmide 24 h (B), 48 h (C) und 72 h (D) nach der Hefetransformation.



**Abbildung III.33: Gelelektrophoretische Auftrennung *Xba*I-verdauter Fusionskonstrukte.** **A:** Plasmide vor der Transformation von *Saccharomyces cerevisiae*; **B:** 24 h nach Transformation aus der Hefekultur reextrahiert; **C:** 48 h nach der Hefetransformation reextrahierte Fusionskonstrukte; **D:** nach 72 h reextrahierte Plasmide; **M:** NEB-Größenstandard; **1-3:** pAG423::T-G-t (erwartetes Restriktionsmuster: 6,2 kb + 3,8 kb + 3,1 kb); **4-6:** pAG423::T-G-t inv (erwartetes Restriktionsmuster: 6,2 kb + 3,8 kb + 3,1 kb); **7-9:** pAG423::T-G-t GA (erwartetes Restriktionsmuster: 6,5 kb + 3,8 kb + 3,1 kb); **10-12:** pAG423::T-G-t P-T (erwartetes Restriktionsmuster: 6,6 kb + 3,8 kb + 3,1 kb); **13-15:** pAG423::T-G-t T-P (erwartetes Restriktionsmuster: 6,6 kb + 3,8 kb + 3,1 kb); **16-18:** pAG423::T-G-t co (erwartetes Restriktionsmuster: 3,8 kb + 3,7 kb + 3,1 kb + 2,7 kb). **Rot:** Markierung für Fragmentierung, die dem erwarteten Restriktionsmuster entspricht.

Ein Vergleich der reextrahierten Plasmid-DNS mit den ursprünglichen Fusionskonstrukten zeigte, dass unabhängig von der Probenahme der insgesamt 18 untersuchten Reextrakte nur B12 (rote Markierung Abbildung III.33) mit dem ursprünglichen Vektor übereinstimmt. Auch wird augenscheinlich, dass keines der Konstrukte innerhalb der Dreifachbestimmung identisch ist oder abhängig von der Inkubationszeit in der Hefezelle eine bestimmte Tendenz hinsichtlich seiner Veränderung zeigt. Um die mittels Restriktion nachgewiesene Abweichung genauer zu analysieren, wurde anhand von Sequenzierungsansätzen die genaue Basenabfolge untersucht. Obwohl das Fragmentierungsmuster bei jedem Konstrukt variierte, wurde die Sequenzierung dabei aufgrund zeitlicher Limitierung auf ein reextrahiertes Fusionsplasmid (pAG423::*T-G-t*, 72 h, Kolonie 2) beschränkt. Abbildung III.34 gibt die erhaltene Basenabfolge im Vergleich mit dem ursprünglichen Konstrukt schematisch wieder, das vollständige Alignment ist im Anhang zu finden.

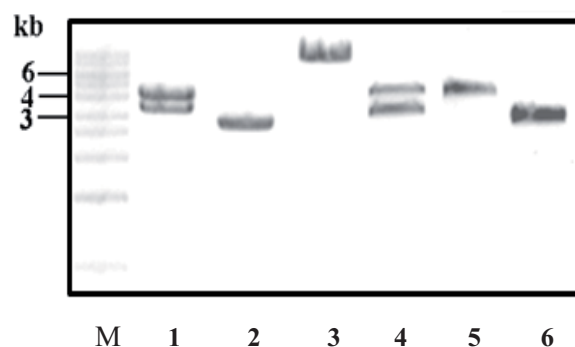


**Abbildung III.34:** Sequenzierungsergebnis des reextrahierten Fusionskonstrukts pAG423::*T-G-t*. Gelber Pfeil: die Basenabfolge des erhaltenen Sequenzierungsergebnisses; gelbe Linien: Ort der entsprechenden Sequenz innerhalb des Fusionskonstrukts vor Transformation von *Saccharomyces cerevisiae*.

Ein Vergleich der Sequenzierungsergebnisse des reextrahierten Fusionskonstrukts mit der ursprünglichen Basenabfolge des in *Saccharomyces cerevisiae* transformierten Plasmids pAG423::*T-G-t* zeigte einen Verlust des kompletten Inserts zwischen dem Zeitpunkt der Transformation der Hefezelle und der Reextraktion des Konstrukts. Die auf dem reextrahierten Plasmid durchgehende Sequenz (gelber Pfeil in Abbildung III.34) entspricht drei getrennten Bereichen des Ursprungsplasmids, von denen der erste eine Länge von 939 bp aufweist und vor

dem vektoreigenen *GPD*-Promotor lokalisiert ist. Die restlichen 145 bp des sequenzierten Fragments teilen sich auf dem in die Transformation eingesetzten Vektor in zwei Abschnitte (48 bp und 77 bp), die beide hinter dem *CYC1*-Terminator liegen und durch 1188 bp voneinander getrennt sind, die den Histidinmarker beinhalten.

Der pAG423-Vektor wurde ursprünglich für die Expression von Einzelgenen konstruiert. Im Rahmen dieser Arbeit wurde hingegen ein Fragment von über 6 kb integriert, welches zusätzlich aus drei Genen und ihren Regulationseinheiten besteht. Um daher zu klären, ob die abweichende Anwendung für die gezeigte Plasmidveränderung verantwortlich ist, wurde pAG423::*T10H* ebenfalls auf eine Veränderung innerhalb der Hefezelle untersucht. Dieser Vektor enthält ausschließlich das 1,5 kb große Gen für die 5 $\alpha$ -Taxadienol-10 $\beta$ -Hydroxylase aus *Taxus cuspidata* (Verweis Einleitung) und war bereits konstruiert worden, um weitere Schritte der Taxolbiosynthese zu untersuchen. Die Sequenzen der für die Klonierung verwendeten Oligonukleotide 47 und 48 sind in Tabelle II.3 angegeben. Gleichzeitig wurde der Leervektor pAG423 im Hinblick auf seine Stabilität innerhalb der Hefezelle analysiert. Wie auch mit den Fusionskonstrukten wurde der *Saccharomyces cerevisiae*-Stamm CEN.PK2-1C mit pAG423 und pAG423::*T10H* transformiert, Plasmid-DNS reextrahiert, über eine Transformation von *E. coli* vereinzelt und anhand einer Restriktion mit dem jeweiligen Ursprungsvektor verglichen. Abbildung III.35 zeigt die gelelektrophoretische Auftrennung der verdauten Plasmide.

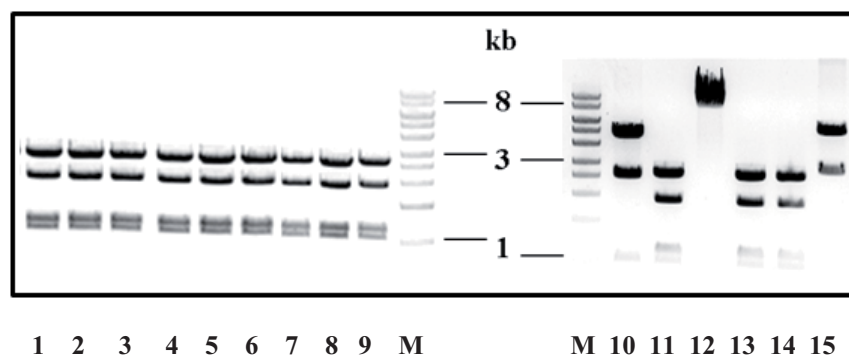


**Abbildung III.35: Gelelektrophoretische Auftrennung des Restriktionsansatzes von Plasmiden vor und nach Transformation der Hefezellen mit *Xba*I.** M: Fermentas-Größenstandard; 1: pAG423::*T10H* vor Transformation von *S. cerevisiae*; 2-3: pAG423::*T10H* nach Transformation von *S. cerevisiae* (erwartetes Fragmentierungsmuster: 3,8 kb + 4,5 kb); 4: pAG423 vor Transformation von *S. cerevisiae*; 5-6: pAG423 nach Transformation von *S. cerevisiae* (erwartetes Fragmentierungsmuster: 3,8 kb + 4,8 kb)



Unabhängig von der Insertion eines Gens wurde die Basenabfolge der untersuchten Plasmide innerhalb der Hefezelle verändert. So weicht sowohl die Fragmentierung des Leervektors pAG423, als auch diejenige von pAG423::*T10H* nach Transformation und Reextraktion von derjenigen des jeweils in die Transformation eingesetzten Konstrukts deutlich ab.

Das Ausmaß der nachgewiesenen Plasmidveränderung der pAG-Vektoren innerhalb des *Saccharomyces cerevisiae* Stamms CEN.PK2-1C steht im Gegensatz zu dem die Einzelvektoren enthaltenen CEN 7. Sowohl der reproduzierbare Nachweis der Expression aller Gene anhand von Northernblots, als auch seine starke und gleichbleibende Produktionsleistung ist mit einer Plasmidveränderung nicht vereinbar. Ein Verlust der Plasmide konnte sowohl bei transformierten pAG-Derivaten, als auch bei den in CEN 7 enthaltenen pVV- und pRS-Vektoren ausgeschlossen werden (III.3.4.1). Allerdings konnte gezeigt werden, dass zumindest die Basenabfolge des kommerziell erhältlichen Vektors pAG423 in der Hefezelle verändert wird. Dieses Rekombinationsereignis erwies sich als unabhängig von der Insertion eines DNS-Fragments, als auch seiner Größe. Um zu untersuchen, ob die für die Generierung von CEN 7 verwendeten pVV- und pRS-Vektoren trotz der stabilen Produktionswerte von Taxadien und GGOH ebenfalls einer Veränderung unterliegen, wurde Gesamtplasmid aus CEN 7 isoliert, durch eine Transformation von *E. coli* vereinzelt und 24 der gewachsenen Kolonien auf ihre Plasmide getestet. In Abbildung III.36 ist die gelelektrophoretische Auftrennung des *Xba*I-Verdaus isolierter Vektoren dargestellt.



**Abbildung III.36: Gelelektrophoretische Auftrennung *Xba*I-verdauter Reextraktionsplasmide 1-15 aus CEN 7.** **M:** Fermentas-Größenstandard. Erwartetes Restriktionsmuster für pVV214::*TDSCO* (4,6kb + 2,5kb + 1kb) 1 mal zu erkennen (10); Erwartetes Restriktionsmuster für pVV200::*GGPPS-Sa* (2,7kb + 1,9kb + 1,1kb + 1kb) 12 mal (1-9,11,13-14); Erwartetes Restriktionsmuster für pRS315::*tHMGR1-Sc* (7,3kb + 742bp + 235bp) 1 mal (12); Erwartetes Restriktionsmuster für pRS313::*UPC2.1-Sc* (5,3kb + 2,8kb + 240bp) 1 mal (15).

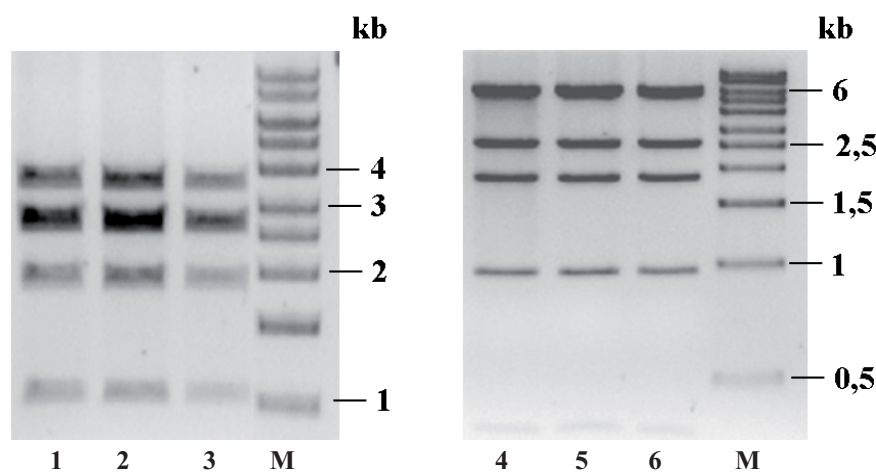


Anhand des *Xba*I-Verdaus von aus CEN 7 reextrahierten Plasmiden konnte gezeigt werden, dass für jeden der vier in den Stamm eingebrachten Vektoren mindestens eins der erhaltenen Fragmentierungsmuster mit dem jeweils virtuell kalkulierten Restriktionsergebnis übereinstimmt. Des Weiteren wurde nach gelelektrophoretischer Auftrennung kein reextrahiertes Plasmid mit abweichender Bandenverteilung beobachtet. Eine Rekombination von pVV- oder pRS-basierten Vektoren konnte somit ausgeschlossen werden. Der jeweilige Anteil von insgesamt 20 reextrahierten Plasmiden variierte dagegen auffällig. Sowohl pVV214::*TDSco*, als auch pRS315::*tHMGR1-Sc* und pRS313::*UPC2.1-Sc* waren jeweils mit 5 % der extrahierten Vektoren vertreten. pVV200::*GGPPS-Sa* dagegen stellte mit 85 % aller Reextraktionen den stabilsten der vier eingebrachten Hybridplasmide dar.

### III.3.5 Klonierung des Genclusters in pVV200

Bei keinem der im Rahmen dieser Arbeit generierten Fusionsstämme konnte ein Transkriptionsnachweis aller Gene des Clusters erbracht werden. Die Reextraktion und Restriktionsanalyse der verschiedenen in den *Saccharomyces cerevisiae*-Stamm CEN.PK2-1C eingebrachten Fusionskonstrukte zeigte dies begründet in einer durch Rekombinationsereignisse innerhalb der Zelle veränderten Basenabfolge aller pAG423-Derivate (III.3.4.2). Die detektierte Veränderung schien weder abhängig zu sein von der Auswahl der Promotor-, bzw. Terminatorsequenzen noch von der Orientierung der einzelnen Gene innerhalb des Clusters zueinander. Sequenzierungsergebnisse eines dieser reextrahierten Vektoren demonstrierte den Verlust des kompletten Genclusters. Durch den Austausch dieses Genclusters gegen ein Einzelgen konnte gezeigt werden, dass die Veränderungen nicht an der ungewöhnlichen Länge des inserierten DNS-Fragments in den pAG-Vektor lag, sondern am Vektorgrundgerüst selbst. Gestützt wurde dies durch die nachgewiesene Rekombination des Leervektors nach Transformation einer Hefekultur. Im Gegensatz dazu konnten alle pVV- und pRS-basierten Hybridplasmide des generierten Stamms CEN 7 unverändert reextrahiert werden (III.3.4.2). Um zu untersuchen, ob diese Vektoren auch bei Expression eines Genclusters funktionell sind, wurde dies in pVV200 inseriert. Dieser Vektor wurde aufgrund seines hohen Anteils an extrahiertem Gesamtplasmid aus CEN 7 (Abbildung III.36) für die Insertion des Genclusters ausgewählt. Aufgrund der zeitlichen Limitierung wurden von den neun generierten Fusionskonstrukten das *TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-tefp-pgkp-tHMGR-Sc1* (Tabelle III.8) für die Umklonierung von pAG423 in pVV200 ausgewählt, da es zur höchsten Taxadienproduktion geführt hatte (Tabelle III.11). Für einen Vergleich mit dem codon-optimierten

Fusionskonstrukt wurde weiterhin *TDSco-80-pho5tt-tefp-GGPPSco-adh1tt-adh1p-tHMGR-Sc1* in pVV200 inseriert. Die Klonierung des Clusters in pVV200 erfolgte mittels Gateway-Technologie. Die generierten Fusionskonstrukte pVV200::*T-G-t* T-P und pVV200::*T-Gco-t* (Tabelle III.8) wurden nach bestätigter korrekter Basenabfolge durch Sequenzierungen verwendet, um CEN.PK2-1C zu transformieren. Analog zu den bereits untersuchten Hefestämmen wurde aus den entstandenen Fusionsstämmen CEN 17-19 (Tabelle III.8) Gesamtplasmid isoliert und einer Testrestriktion mit der Endonuklease *XbaI* unterzogen. Abbildung III.37 zeigt die gelelektrophoretische Auftrennung der verdauten reextrahierten Plasmide, sowie der Fusionsplasmide vor der Transformation von CEN.PK2-1C.



**Abbildung III.37: Gelelektrophoretische Auftrennung der Testrestriktion reextrahierter auf pVV200 basierender Fusionskonstrukte. 1-3:** *XbaI*-verdaute Reextraktionsplasmide pVV200::*T-Gco-t* aus CEN 17 (erwartetes Fragmentierungsmuster: 3,7 kb + 2,7 kb (2x) + 1,9 kb + 1 kb); **M:** Fermentas-Größenstandard; **4-6:** *XbaI*-verdaute Reextraktionsplasmide pVV200::*T-G-t* T-P aus CEN 18 (erwartetes Fragmentierungsmuster: 6,6 kb + 2,7 kb + 1,9 kb + 1 kb + 350 bp).

Alle reextrahierten Fusionskonstrukte zeigten bei Verdau mit *XbaI* ein Fragmentierungsmuster, das mit dem jeweiligen ursprünglichen Fusionsvektor übereinstimmte. So waren bei Verdau reextrahierter pVV200::*T-Gco-t* aus CEN 17 Banden bei 3,7 kb + 2,7 kb (2x) + 1,9 kb + 1,0 kb zu detektieren, während bei pVV200::*T-G-t* T-P aus CEN 18 nach Verdau Fragmente der Größe 6,6 kb + 2,7 kb + 1,9 kb + 1 kb + 350 bp nachzuweisen waren. Eine Veränderung des Vektors pVV200 konnte demnach auch bei Insertion verschiedener Gencluster nicht detektiert werden.

Mittels GC/MS-Messungen wurden CEN 17 und 18 auf die jeweilige Produktion von Taxadien, GGOH und Squalen hin untersucht. Die Quantifizierung der detektierten Mengen ist in Tabelle III.13 angegeben. Als Vergleichsstämme wurden CEN 12 und CEN 16 aufgeführt, in die die

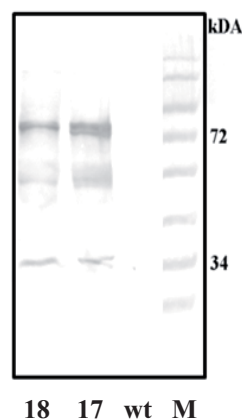
jeweiligen Konstrukte im pAG-Vektor transformiert worden war. Der Wildtypstamm bildete die Negativ-, CEN 7 mit den in Einzelvektoren exprimierten Genen die Positivkontrolle.

**Tabelle III.13: Quantifizierung selektiver Produkte einzelner Fusionsstämmen via GC/MS**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>G</i> ; <i>t</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±3,2	3,90±0,3
CEN 16	pAG:: <i>T</i> -pho5tt-tefp- <i>Gco</i> -adh1tt-adh1p- <i>t</i>	0,46±0,005	0,36±0,007	4,00±0,1
CEN 17	pVV:: <i>T</i> -pho5tt-tefp- <i>Gco</i> -adh1tt-adh1p- <i>t</i>	2,70±0,1	15,4±0,3	31,3±3,2
CEN 12	pAG:: <i>T</i> -pho5tt-gcn4tt- <i>G</i> -tefp-pgkp- <i>t</i>	1,86±0,1	2,50±0,1	26,5±7
CEN 18	pVV:: <i>T</i> -pho5tt-gcn4tt- <i>G</i> -tefp-pgkp- <i>t</i>	3,00±0,1	12,4±0,3	60,3±5,2

**T:** *TDSco*-80; **G:** *GGPPS*-Sa; **t:** *tHMGR*-Sc; **pho5TT:** Terminator der sauren Phosphatase **tefp:** Promotor des translationalen Elongationsfaktorgens; **Gco:** für *S. cerevisiae* codon-optimiertes Geranylgeranyldiphosphat-Synthasegen aus *Sulfolobus acidocaldarius*; **adh1tt:** Terminatorregion des Alkoholdehydrogenasegens; **adh1P:** Alkoholdehydrogenase-Promotor;

Ein Vergleich der extrahierbaren Mengen an Taxadien zwischen CEN 16 und CEN 17 zeigte mit 2,69 mg/l eine sechsfache Steigerung bei Expression des Genclusters im pVV200-Vektor gegenüber derjenigen im pAG423. Auch die detektierbare GGOH-Konzentration von CEN 17 war mit 15,4 mg/l 43 mal höher als in CEN 16. Des Weiteren konnte aus CEN 17 eine Squalenmenge extrahiert werden, die die wildtypähnliche Konzentration von CEN 16 um fast das achtfache überstieg. Obwohl die Umklonierung des Konstrukts *T-G-t* T-P vom pAG423-Vektor in den pVV200-Vektor zu einer verdoppelten Synthese des Nebenprodukts Squalen führte, konnte gleichzeitig eine 60%ige Taxadienproduktion und eine fünffache Menge an gebildetem GGOH beobachtet werden. Die Untersuchung der Proteinfraktion von CEN 17 und 18 zeigt jeweils eine Bande richtiger Größe für *TDSco*, *GGPPS*, bzw. *GGPPS*-Sa und *tHMGR*-Sc und somit die Akkumulation aller rekombinanten Enzyme (Abbildung III.38).



**Abbildung III.38: Westernblot von 40 µg TCA-gefällten Gesamtproteins der generierten Fusionsstämmen CEN 17 und 18.** In beiden Stämmen werden die jeweiligen Cluster über pVV200 exprimiert **18:** CEN 18 (Fusionskonstrukt: *TDSco*-80-pho5tt-gcn4tt-*GGPPS*-Sa-tefp-pgkp-*tHMGR*1); **17:** CEN 17 (Fusionskonstrukt: *TDSco*-80-pho5tt-tefp-*GGPPS*-Sa-adh1tt-adh1p-*tHMGR*1); **Wt:** CENPK2-1C Wildtyp; **M:** Fermentas-Größenstandard

### III.4 Konstruktion eines künstlichen Hefechromosoms

Die in III.3 generierten Fusionsvektoren basieren alle auf dem Vektorgrundgerüst von Plasmiden, die ursprünglich für die Expression eines Einzelgens konstruiert wurden. Obwohl die Insertion dreier Gene mitsamt Promotor und Terminator zu einem stabilen Vektor und erfolgreicher Expression führte (III.3.5), ist eine Größenlimitierung nicht auszuschließen. Künstliche Hefechromosomen dagegen werden nicht nur speziell für große DNS-Fragmente verwendet, sondern zeichnen sich des Weiteren durch eine hohe Stabilität innerhalb der Hefezelle aus.<sup>[158]</sup> Die Umklonierung des generierten Genclusters vom Plasmidgrundgerüst in ein YAC stellt daher eine geeignete Plattform für die weitere Insertion von Genen aus der Taxolbiosynthese dar. Für diesen Ansatz wurde das von Burke et al 1987 generierte pYAC4 ausgewählt. Es enthält sowohl die benötigten Sequenzen für die Replikation und den Erhalt in *E. coli*-Zellen (pMB1-ori und Ampicillinresistenzgen), als auch für die Selektion in *Saccharomyces cerevisiae* (Tryptophan-, Histidin-, Leucinmarker) sowie die dortige, an den Zellzyklus gebundene Replikation (ars1) und Segregation (CEN 4). Für die Stabilität als lineares Chromosom wurden Telomersequenzen eingefügt. Das pYAC4 wurde bereits erfolgreich zur Kartierung von Genomen und Erstellung von Genombibliotheken verwendet,<sup>[227-230]</sup> sowie zur Anfertigung einer Bibliothek verschiedener Konstellationen von sieben Genen aus der Flavonoid-Biosynthese.<sup>[231]</sup> Neben seiner eigentlichen Anwendung als linearisiertes künstliches Hefechromosom ist es auch in zirkulärer Form einsetzbar, was eine Isolierung des generierten YACs und die Sequenzierung des Inserts erleichtert.<sup>[232]</sup>

#### III.4.1 Modifikation von pYAC4

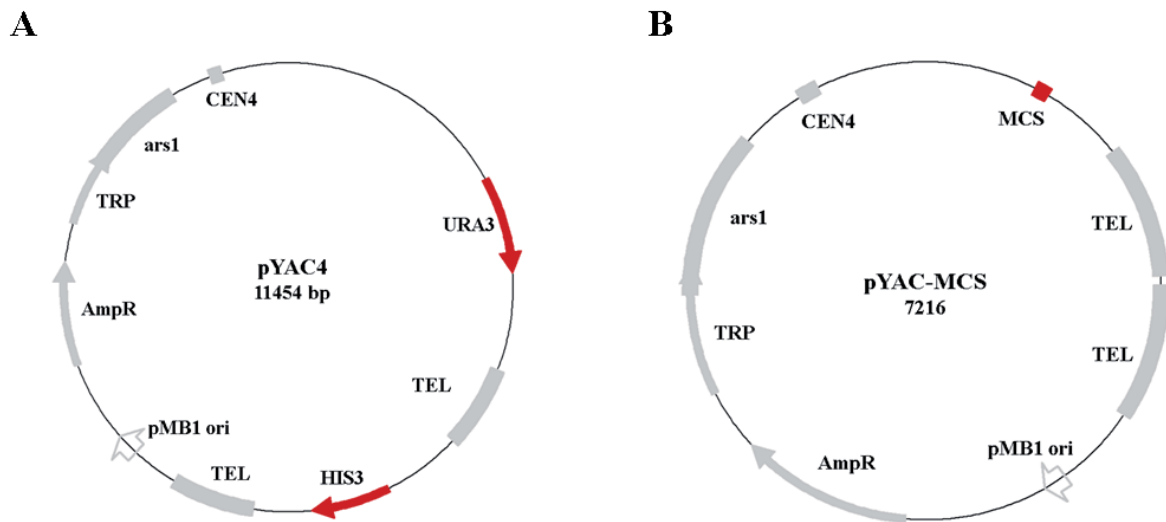
Für die Insertion von DNS-Abschnitten in pYAC4 wird dieses mit Restriktionsendonukleasen verdaut, sodass zwei Fragmente entstehen, die jeweils an einer Seite von Telomersequenzen begrenzt sind. Diese Fragmente werden anschließend mit ebenfalls verdauter DNS ligiert, wodurch ein Konstrukt entsteht, welches aus der zu analysierenden DNS besteht, die von den beiden Armen des Chromosoms flankiert ist. Die Selektion korrekter Liganden erfolgt nach Transformation in Hefespheroblasten anhand des Selektionsmarkers in jedem der beiden Teilabschnitte des YAC-Grundgerüsts. Die Anzahl an inserierten DNS-Fragmenten kann dabei variiert werden. Naesby und Mitarbeiter konnten durch die Transformation aller Gene des Flavonol-Biosynthesewegs in *Saccharomyces cerevisiae*-Spheroblasten einen Zusammenbau von bis zu 50 Expressionskassetten erzielen.<sup>[231]</sup>

Im Rahmen dieser Arbeit wurden die zu exprimierenden Gene bereits zu einem Cluster fusioniert, sodass ein Zusammenbau mehrerer DNS-Abschnitte nicht benötigt wurde. Um eventuelle Modifikationen oder Erweiterungen zu erleichtern, sowie einen direkten Vergleich zwischen dem pVV200-basierten mit dem YAC-basierten Fusionskonstrukt durchzuführen, sollte das Gencluster des Weiteren zunächst in der zirkulären Form des YACs exprimiert werden. Anstelle einer Trennung der YAC-Arme mit anschließender Ligation der fusionierten Gene mit den beiden YAC-Fragmenten wurde das Cluster daher in das linearisierte Plasmid integriert.

Aufgrund dieses abweichenden Klonierungsansatzes bestand das ursprüngliche pYAC4 neben den beschriebenen auch aus Sequenzabschnitten, die weder für die Expression des Genclusters noch für eventuelle Modifikationen benötigt werden. Gleichzeitig fehlte es an singulären Restriktionsschnittstellen, die weder im Gencluster noch in weiteren Genen der Taxolbiosynthese vorkommen. Derartige Schnittstellen sind nicht nur für die Integration elementar, sondern auch für die Möglichkeit einer finalen Linearisierung zur Freilegung der Telomere. Um das pYAC4-Grundgerüst in die gewünschte Form zu bringen, wurden vor Integration des Genclusters daher folgende Modifizierungen vorgenommen:

1. *Entfernen des nicht benötigten Histidin-Markers zwischen den Telomersequenzen bei gleichzeitiger Integration einer PmeI-Schnittstelle.* Damit das zirkuläre pYAC als künstliches Chromosom eingesetzt werden kann, muss es so linearisiert werden, dass die Telomersequenzen endständig lokalisiert sind. Dazu ist eine Restriktion mit *Bam*HI notwendig. Diese Schnittstelle befindet sich allerdings mehrfach in dem integrierten Gencluster und kann demzufolge zur finalen Linearisierung des YACs nicht verwendet werden. Der Histidinmarker wurde daher durch einen synthetisierten Linker mit der singulären *PmeI*- Erkennungssequenz ersetzt.
2. *Austausch des nicht benötigten Uracil-Selektionsmarkers durch eine Multiple cloning site MCS.* Um nicht nur das generierte Gencluster, sondern eventuell auch zusätzliche Gene der Taxolbiosynthese in pYAC integrieren zu können, wurden Erkennungssequenzen für Restriktionsendonukleasen benötigt, die in dem Konstrukt selbst nicht schneiden. Daher wurde die nicht verwendete Sequenz für das Orotidin 5-phosphat-Decarboxylasegen, welches die Uracil-Auxotrophie komplementiert, durch eine MCS mit mehreren einmaligen Restriktionsschnittstellen ersetzt.

Das ursprüngliche, sowie das für die Expression des generierten Genclusters modifizierte pYAC4 sind in Abbildung III.39 dargestellt.



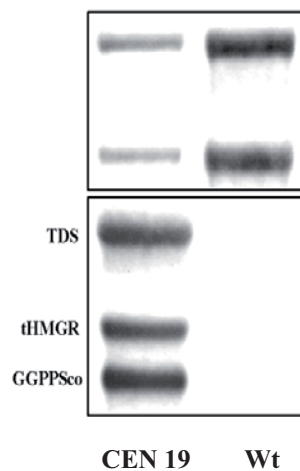
**Abbildung III.39: Modifizierung von pYAC4 für die anschließende Insertion des Genclusters. A:** Ausgangsplasmid pYAC4; **B:** Im Rahmen dieser Arbeit modifiziertes pYAC4 für Generierung eines Fusionschromosoms. **Rot:** modifizierte Bereiche; **CEN4:** Centromer von Chromosom 4 aus *S. cerevisiae*; **URA3:** Uracil-Marker (Orotidin 5-phosphat Decarboxylasegen); **TEL:** *Tetrahymena*-Telomer; **HIS3:** Histidin-Marker (Imidazoglycerolphosphat Dehydrogenasegen); **pMB1 ori:** bakterieller Replikationsursprung; **AmpR:** Ampicillin-Resistenzgen (Beta-Laktamase); **TRP1:** Tryptophan-Marker (Imidazoglycerolphosphat Dehydratasegen); **ars1:** Hefereplikator; **MCS:** *Multiple cloning site*

#### III.4.2 Integration des Genclusters und Generierung eines rekombinanten Hefestamms

Grundsätzlich zeigten die Fusionsstämme CEN 17 und 18 ähnliche Mengen an Taxadien und GGOH (Tabelle III.13). Da die Expression des Genclusters *T-Gco-t* verglichen mit *T-G-t* T-P nur zu der Hälfte der Menge an dem Nebenprodukt Squalen führte, wurde dieses für die Umklonierung in das modifizierte YAC ausgewählt. Die fusionierten Gene wurden dazu inklusive des vektoreigenen *GPD*-Promotors und *CYC1*-Terminators aus dem Fusionsvektor amplifiziert (Primer 51 und 52, Tabelle II.3). Das erhaltene PCR-Produkt wurde anschließend anhand der angefügten *NotI*-Erkennungssequenz verdaut und in das ebenfalls durch Inkubation mit der Restriktionsendonuklease *NotI* linearisierte pYAC ligiert. Nach einer Verifizierung der Insertion durch eine Testrestriktion mit *XbaI* wurde die korrekte Basenabfolge des generierten Fusionskonstrukts durch Sequenzierungen bestätigt. Durch Transformation des *S. cerevisiae*-Stamms CEN.PK2-1C mit pYAC::*T-Gco-t* wurde der Stamm CEN 19 generiert.



Anhand eines Northernblots konnte die Expression der inserierten Gene nachgewiesen werden (Abbildung III.40).



**Abbildung III.40: Northernblotanalyse des im YAC exprimierten Fusionskonstrukts. Oben:** Gelelektrophoretisch aufgetrennte Gesamt-RNS, jeweils 10 µg geladen. **Unten:** Autoradiogramm der spezifisch gebundenen radioaktiv markierten DNS-Sonde; **CEN 19:** Fusionskonstrukt *T-Gco-t* mithilfe eines YACs exprimiert; **Wt:** CENPK2-1C Wildtyp

Um den Stamm auf seine Produktion an Taxadien, GGOH und Squalen zu untersuchen, wurde er wie in II.1.4.2 beschrieben angezogen, mit Cyclohexan extrahiert und die jeweiligen Mengen mittels GC/MS detektiert und quantifiziert. Tabelle III.14 gibt die ermittelten Werte wieder.

**Tabelle III.14: Produktion von Taxadien, GGOH und Squalen verschiedener Fusionsstämme**

Stamm	Überexprimierte Gene	Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/l]
wt	keine	n.d.	0,23±0,003	3,10±0,1
CEN 7	<i>TDSco</i> ; <i>G</i> ; <i>t</i> ; <i>UPC2.1-Sc</i>	8,70±1	33,1±3,2	3,90±0,3
CEN 16	<i>pAG::T-pho5tt-tefp-Gco-adh1tt-adh1p-t</i>	0,45±0,005	0,36±0,007	4,00±0,1
CEN 17	<i>pVV::T-pho5tt-tefp-Gco-adh1tt-adh1p-t</i>	2,70±0,1	15,4±0,3	31,3±3,2
CEN 19	<i>pYAC4::gdp- T-pho5tt-tefp-Gco-adh1tt-adh1p-t-cyc1tt</i>	0,56±0,015	3,90±0,3	53,8±5,2

**T:** *TDSco-80*; **G:** *GGPPS-Sa*; **t:** *tHMGR-Sc*; **pho5TT:** Terminator der sauren Phosphatase; **tefP:** Promotor des translationalen Elongationsfaktorgens; **Gco:** für die Expression in *S. cerevisiae* codon-optimiertes Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **adh1tt:** Terminatorregion des Alkoholdehydrogenasegens; **adh1P:** Alkoholdehydrogenase-Promotor; **gdpP:** Glycerinaldehyd-3-Phosphat-Dehydrogenase-Promotor; **cyc1TT:** Cytochrom1-Terminator

Die Expression der drei Gene *TDSco*, *GGPPco* und *tHMGR1-Sc* mittels künstlichen Hefechromosoms führt zu einer Produktion von Taxadien und GGOH, die nur ein Viertel derjenigen beträgt, die auf eine Expression der Gene über den *high-copy*-Vektor pVV200 zurückzuführen ist. Im Gegensatz dazu verdoppelt sich die Squalenmenge nahezu. Im Vergleich zu CEN 16, der die Gene über den pAG423-Vektor exprimiert, verzehnfachen sich die GGOH- und Squalenmengen bei Expression mit dem YAC, während die Taxadienausbeute von 0,46 mg/l auf 0,56 mg/l nur leicht ansteigt.



## IV Diskussion

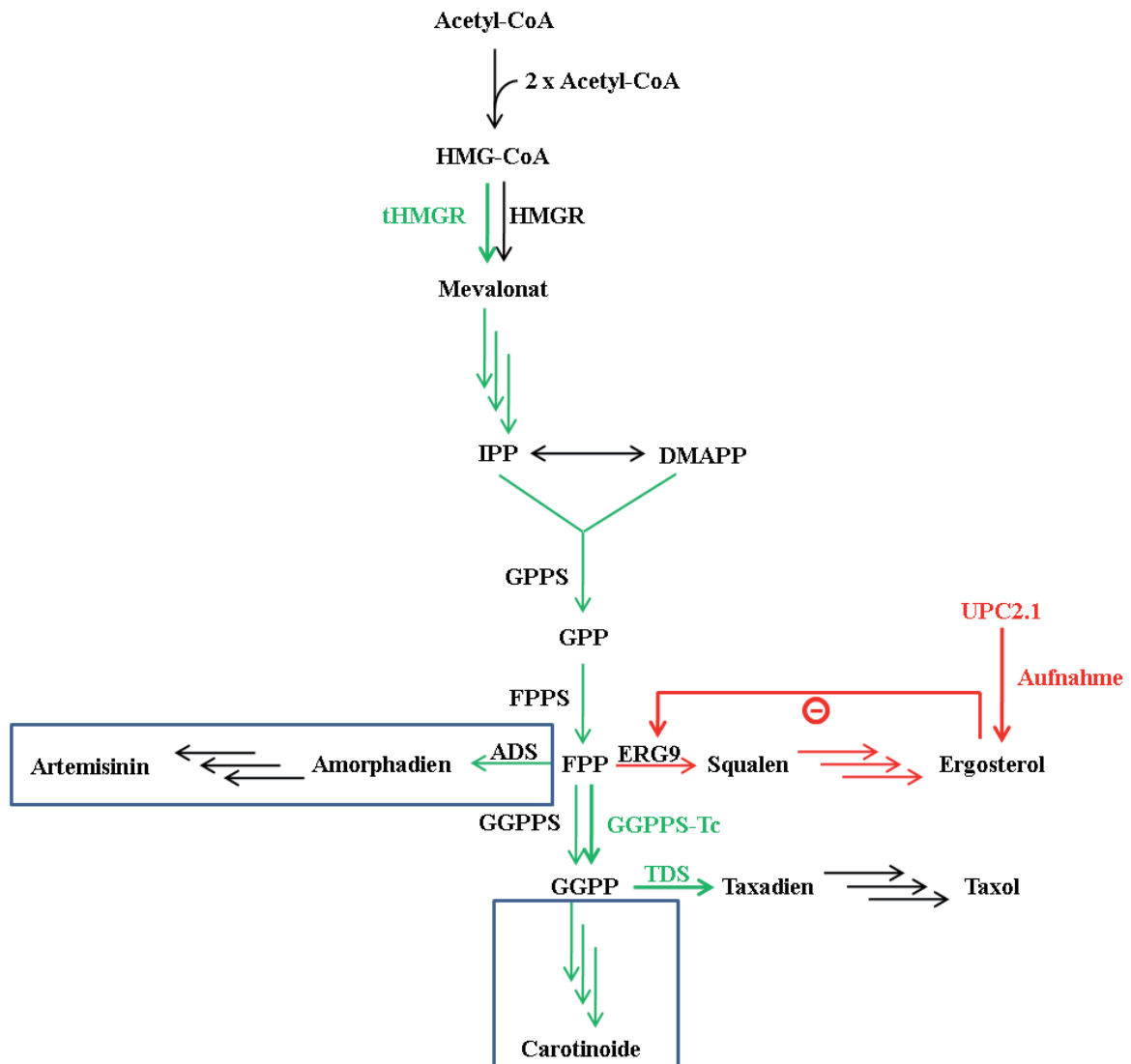
### IV.1 Generierung eines Taxadien produzierenden Hefestammes

Für die rekombinante Produktion von Taxadien in *S. cerevisiae* wurde der Hefestamm CEN.PK2-1C nicht nur um das Taxadiensynthasegen aus der pflanzlichen Taxolbiosynthese erweitert, sondern auch in seinem Primärmetabolismus verändert. Die Notwendigkeit einer Modifikation des Primärstoffwechsels ergibt sich aus mehreren Literaturquellen, in denen bereits gezeigt wurde, dass die Ergänzung des hefeeigenen Metabolismus um die entsprechenden Enzyme eines fremden Diterpenstoffwechselwegs allein zu keiner erfolgreichen Synthese der jeweils gewünschten terpenoiden Verbindung führt.<sup>[124][149][233]</sup> Im Gegensatz zu FPP, welches unter anderem für die Synthese von Ergosterol eingesetzt wird, ist GGPP für das vegetative Wachstum von *S. cerevisiae* nicht notwendig.<sup>[234]</sup> Die niedrige katalytische Aktivität der hefeeigenen GGPP-Synthase führt zu keiner ausreichenden Bereitstellung von GGPP, von der wiederum die weitere Verstoffwechselung zu rekombinant produzierten Diterpenen abhängt. Nicht nur DeJong verwendete daher für die effiziente Produktion von Taxadien neben der Taxadiensynthase gleichzeitig auch die pflanzliche GGPP-Synthase.<sup>[149]</sup> Ebenso exprimierte Yamano für eine *de novo* Synthese von Carotinoiden die benötigten Biosynthesegene für die Umwandlung von FPP über GGPP bis hin zum Lykopen.<sup>[124]</sup> Ohne zusätzliche GGPPS konnten Ukibe *et al.* zwar eine geringe Akkumulation von  $\beta$ -Carotin beobachten, erzielten aber durch die Expression eines weiteren GGPPS-Gens eine 20 fach höhere Ausbeute.<sup>[235]</sup> Es ist daher nach dem Stand der Literatur plausibel, dass auch im Rahmen dieser Arbeit erst die gleichzeitige Expression von TDS- und GGPPS-Gen zu einer quantifizierbaren Taxadienproduktion führte.

Es ist hinreichend dokumentiert, dass eine effiziente Bereitstellung der isoprenoiden Vorstufen IPP und DMAPP absolut notwendig ist, wenn diese zu terpenoiden Verbindungen verknüpft werden sollen. Um dies zu erreichen, wurden sowohl für Prokaryoten als auch für Pflanzen und Hefen diverse Ansätze zur Modifizierung ihres Primärstoffwechsels entwickelt. Alle basieren auf der Umgehung von Limitierungen des Mevalonat- oder DXP-Wegs. So wurden bei der Synthese der Vorstufen über den DXP-Weg entweder entsprechende Gene des Wegs selbst überexprimiert<sup>[131][133][142][143][236-238]</sup> oder die Menge der Ausgangssubstanzen aus dem zentralen Stoffwechselweg erhöht.<sup>[136][137][139-141]</sup> Eine ähnliche Vorgehensweise wurde auch zur Modifizierung des zentralen Hefestoffwechsels<sup>[239]</sup> oder zur Optimierung des Mevalonatwegs basierend auf der Expression des verkürzten HMG-CoA-Reduktasegens *tHMGR* gewählt.<sup>[125][233][240][241]</sup> Der um die *tHMGR* erweiterte Hefestamm CEN 3 dieser Arbeit zeigte

ebenfalls eine Taxadiensteigerung im Vergleich zu CEN 2 ohne entsprechende Expression von *tHMGR*. Allerdings ist die dadurch um 50 % erhöhte Ausbeute an Taxadien nicht mit Literaturwerten zur rekombinanten Produktion anderer Terpene zu vergleichen. Dort wurde die Produktmenge allein durch *tHMGR* zwischen 8-fach<sup>[69]</sup> und 40-fach<sup>[242]</sup> gesteigert. Im Rahmen dieser Arbeit wurde *tHMGR* im Gegensatz zu *TDS* und *GGPPS* nicht über ein *high-copy* Plasmid exprimiert, sondern in einen pRS-Vektor integriert, der einfach pro Zelle vorliegt, sodass die niedrige Kopienzahl des Gens für die geringe Produktivitätssteigerung entscheidend sein könnte. Allerdings wurde bereits gezeigt, dass der Effekt von *tHMGR* auf die Akkumulation isoprenoider Produkte nicht abhängig von der synthetisierten Reduktasemenge oder -aktivität ist, sondern eher auf der höheren Stabilität des verkürzten Enzyms basiert.<sup>[242]</sup> So wurde bei Expression eines chromosomal integrierten *tHMGR* zwar nur eine dreifache Aktivitätssteigerung der Reduktase, bei Expression des gleichen Gens mithilfe eines episomalen *high-copy* Vektors hingegen eine 44-fache Steigerung beobachtet. Trotz dieses Unterschieds zeigten beide Stämme aber die gleiche, 40-fach erhöhte Produktakkumulation.<sup>[242]</sup> Die vergleichsweise geringe Taxadienproduktion trotz Optimierung des Primärmetabolismus in dieser Studie lässt sich daher nicht alleine durch die Wahl des Expressionsvektors für *tHMGR1* mit dessen niedriger Kopienzahl erklären. Auch ist sie nicht in einer ineffizienten Transkription von *tHMGR* oder Translation der entsprechenden mRNS begründet, da der jeweilige positive Nachweis anhand von Northern- (Abbildung III.2), bzw. Westernblotanalysen (Abbildung III.4) erbracht werden konnte. Die zusätzliche Expression von *upc2.1* führte im Rahmen dieser Arbeit ebenfalls nicht zu der von Jackson et al. beobachteten Steigerung der Terpenproduktion von 40%.<sup>[126]</sup> Trotz nachweislich transkribiertem *upc2.1* (Abbildung III.5) zeigte CEN 4 keine erhöhte Taxadienakkumulation.

Bisherige Studien zur rekombinanten Produktion von Taxadien involvierten weder die Expression von *tHMGR* noch von *upc2.1*,<sup>[149]</sup> sodass der Effekt ihrer Genprodukte auf die Synthese von Taxadien selbst noch nicht untersucht wurde. Allerdings konnte bereits gezeigt werden, dass die Überexpression von *tHMGR* und *upc2.1* sowohl zu einer gesteigerten Synthese FPP-abgeleiteter Verbindungen wie Squalen oder Amorphadien, als auch rekombinanter GGPP-abgeleiteter Verbindungen wie Lycopin, Phytoen oder  $\beta$ -Carotin führte.<sup>[125][233][240][241]</sup> Die durch *tHMGR* und *UPC2.1* verstärkt synthetisierten IPP und DMAPP werden dabei zu GPP verknüpft und durch Addition eines weiteren IPP-Moleküls wird FPP synthetisiert (Abbildung IV.1). FPP kann anschließend sowohl für die Produktion sesqui- oder triterpenoider Verbindungen verwendet, als auch durch die Weiterreaktion mit IPP zu GGPP umgesetzt werden und verkörpert dadurch den Knotenpunkt zweier um die isoprenoiden Vorstufen konkurrierenden Biosynthesewege.



**Abbildung IV.1: Schematische Darstellung modifizierter Isoprenoidsynthese-Netzwerke in *S. cerevisiae*.** **Fett + grün:** rekombinante Proteine; **Grün:** nachfolgende Syntheserouten, die durch rekombinante Proteine ebenfalls verstärkt werden. **Rot:** durch UPC2.1 (und die dadurch ermöglichte Sterolaufnahme) verminderte Syntheseroute von FPP zu Sterolen. **Blau:** Markierung von rekombinanten Syntheserouten aus anderen Studien.<sup>[240][233][125][241]</sup> **HMGR:** HMG-CoA-Reduktase aus *S. cerevisiae*; **tHMGR:** um regulative Domäne verkürzte HMGR; **IPP:** Isopentenylidiphosphat; **DMAPP:** Dimethylallyldiphosphat; **GPP:** Geranylidiphosphat; **GPPS:** GPP-Synthase; **FPP:** Farnesylidiphosphat; **FPPS:** FPP-Synthase; **ADS:** Amorphadiensynthase; **ERG9:** Squalensynthase; **GGPP:** Geranylgeranylidiphosphat; **GGPPS:** GGPP-Synthase aus *S. cerevisiae*; **GGPPS-Tc:** GGPP-Synthase aus *T. chinensis*; **TDS:** Taxadiensynthase

Im Rahmen dieser Arbeit war weder eine signifikante Produktivitätssteigerung des Diterpens Taxadien selbst zu detektieren, noch von GGOH als GC/MS-gängigem Hydrolyseprodukt der Diterpenvorstufe GGPP. Daher wurde der Einbau synthetisierten IPPs und DMAPPs in FPP-abgeleitete Folgeprodukte untersucht. Ergosterol ist das Hauptsterol der Hefe und leitet sich von Farnesylidiphosphat FPP ab, welches mit einem zweiten FPP-Molekül zu Squalen, der ersten

spezifischen Vorstufe des Ergosterols, kondensiert wird. In *S. cerevisiae* wird weiterhin der größte Teil an IPP und DMAPP für die Biosynthese von Ergosterol eingesetzt, da diese Verbindung für die Stabilität und Fluidität der Zellmembran sorgt.<sup>[191-193]</sup> Die Analyse des generierten Stamms auf eine Verwendung synthetisierter Isoprenoidvorstufen in FPP-abgeleitete Verbindungen erfolgte daher anhand einer Quantifizierung der Squalenakkumulation. Diese ergab verglichen mit dem Wildtypstamm einen fünffach höheren Wert in CEN 4 (Tabelle IV.1).

**Tabelle IV.1: Quantifizierung synthetisierten Squalens mittels GC/MS**

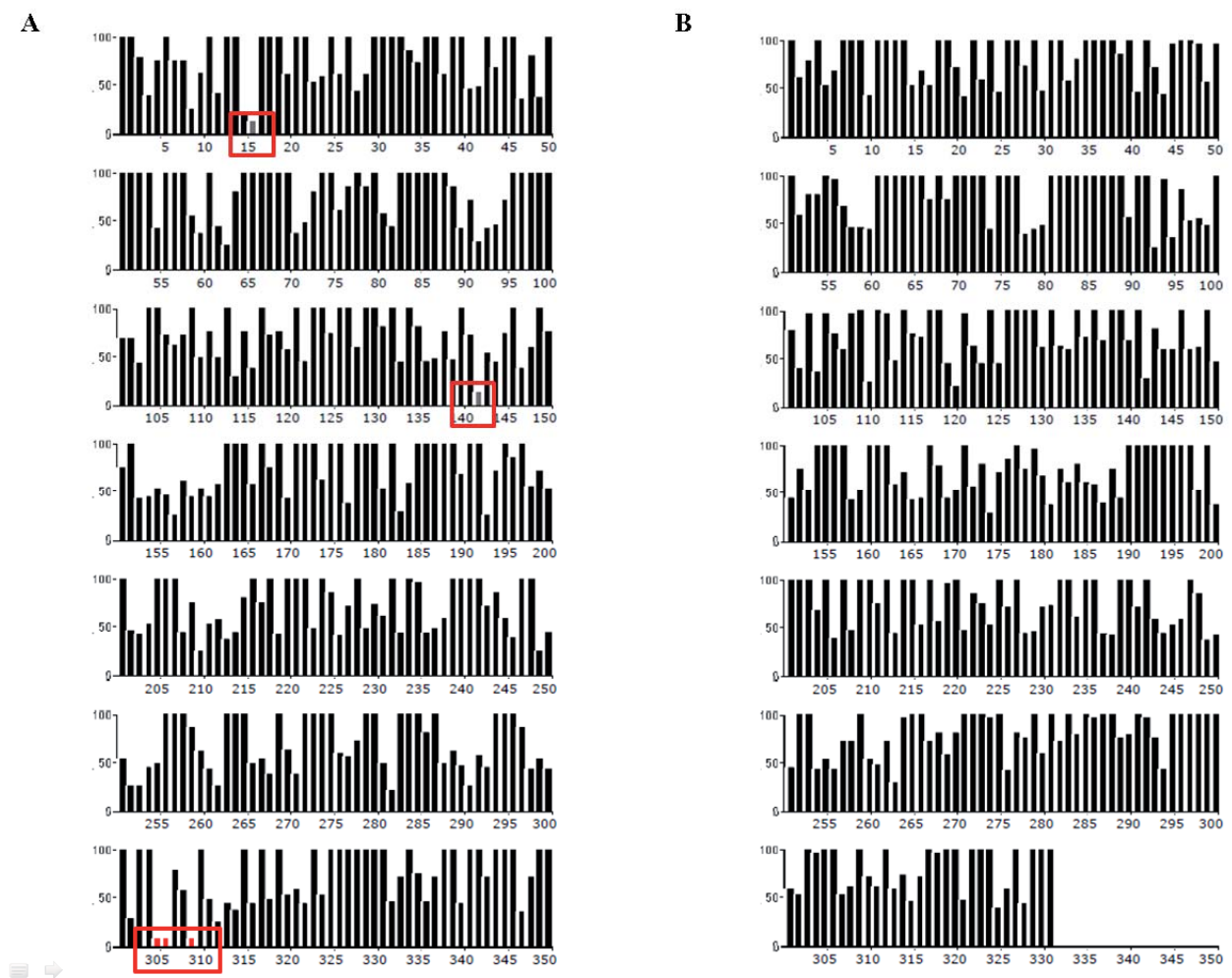
Stamm	Überexprimierte Gene	GGOH [mg/L]	Squalen [mg/L]
CEN wt	keine	0,23±0,003	3,1±0,1
CEN 4	<i>TDS-Tc</i> ; <i>GGPPS-Tc</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	0,28±0,032	16,4±0,2
CEN 5	<i>TDS-Tc</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR1-Sc</i> ; <i>UPC2.1-Sc</i>	27,6±3,4	4,9±0,3

**Wt:** CEN.PK2-1C Wildtyp; ***TDS-Tc*:** Taxadiensynthasegen aus *Taxus chinensis*; ***GGPPS-Tc*:** Geranylgeranyldiphosphatsynthase aus *Taxus chinensis*; ***GGPPS-Sa*:** Geranylgeranyldiphosphatsynthase aus *Sulfolobus acidocaldarius*; ***tHMGR1-Sc*:** HMG-CoA-Reduktase Isozym 1 aus *S. cerevisiae*; ***UPC2.1-Sc*:** Punktmutierte Version des Transkriptionsfaktors UPC2 aus *S. cerevisiae*

Die verstärkte Nutzung der gebildeten Isoprenoidvorstufen IPP und DMAPP für die Synthese von Squalen und weiter von Ergosterol wurde durch die Expression von *GGPPS* aus *S. acidocaldarius* (*GGPPS-Sa*) umgangen. Dieses Gen kodiert für ein bifunktionales Enzym, welches nicht nur die Reaktion von FPP mit IPP zu GGPP, sondern alle Kondensationsschritte ausgehend von IPP und DMAPP bis hin zu GGPP kondensiert.<sup>[210a]</sup> Dadurch wird das produzierte Intermediat FPP von demselben Enzym, welches bereits seine Synthese katalysiert hat, direkt weiter zu GGPP umgesetzt und steht daher für die Ergosterolbiosynthese nicht zur Verfügung. Bei Austausch des pflanzlichen durch das archaebakterielle GGPP-Synthasegen sank die Squalenakkumulation daher von 16,4 mg/L auf 4,9 mg/L (Tabelle IV.1).

Die durch den Austausch des GGPP-Synthasegens bedingte hundertfache Steigerung der GGPP-Produktion basiert allerdings nicht nur auf den bifunktionalen Eigenschaften des archaebakteriellen Enzyms als gleichzeitige FPP- und GGPP-Synthase. Obwohl beide Enzyme nachweislich His-markiert waren, konnte ausschließlich GGPPS-Sa auf Proteinebene detektiert werden (Abbildung III.4). Die Aktivität der pflanzlichen GGPP-Synthase konnte dagegen nur indirekt anhand einer erhöhten GGOH- und Taxadienproduktion nachgewiesen werden (Tabelle III.3).

Ein Vergleich des pflanzlichen Codongebrauchs mit demjenigen des Expressionsstamms *S. cerevisiae* (Abbildung IV.2) weist Mangelcodons an fünf Stellen auf, die eine problematische Translation der nachweislich gebildeten mRNA (Abbildung III.5) bedingen können.



**Abbildung IV.2:** Vergleich der jeweiligen *Codon-usage* von GGPP-Synthase aus *Taxus chinensis* (A) und *Sulfolobus acidocaldarius* (B) mit derer von *S. cerevisiae* (codon usage graphical analyser 2.0). Angegeben in relativer Anpassung an die Codonverteilung in %. **Grau:** unter 20 %; **Rot:** unter 10 %.

Die durchschnittliche Abweichung des archaebakteriellen Codongebrauchs von demjenigen in *S. cerevisiae* ist mit 21,84% nur geringfügig niedriger als diejenige des pflanzlichen Gens mit 25,89% ([www.gcua.de](http://www.gcua.de)). Jedoch sind darin keine Mangelcodons wie beim pflanzlichen Gen enthalten, was zu einer effizienteren Translation und dementsprechend größeren Proteinakkumulation führt (Abbildung III.4). Bedingt durch die unterschiedlichen Mengen an gebildetem Enzym kann allerdings nicht eindeutig geklärt werden, ob die erhöhte GGOH-Produktion von CEN 5 im Gegensatz zu CEN 4 auf einer höheren Aktivität der archaebakteriellen GGPPS-Sa im Vergleich zur pflanzlichen GGPP-Tc, bzw. auf seiner Bifunktionalität beruht oder ausschließlich auf der größeren Enzymmenge. Es gibt bisher keine publizierten Studien mit kinetischen Daten aufgereinigter GGPPS-Sa. Einen Anhaltspunkt bietet allenfalls der Vergleich

partiell gereinigter GGPP-Synthasen. Mit  $8,7 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  Protein zeigt die angereinigte archaebakterielle GGPPS<sup>[210a]</sup> eine 2 mal höhere spezifische Aktivität als eine angereinigte GGPPS aus *Taxus* ( $3,8 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  Protein,<sup>[243]</sup>). Eine gesteigerte GGOH-Produktion durch die Expression des archaebakteriellen GGPP-Synthasegens wäre daher auch im Vergleich mit einer codonoptimierten Version des pflanzlichen Gens plausibel.

Ähnlich der Expression des pflanzlichen *GGPPS-Tc* konnte auch bei *TDS-Tc* trotz detekrierbarer Taxadienproduktion (Tabelle III.4) anhand einer Westernblotanalyse kein Protein nachgewiesen werden (Abbildung III.6). Ein Vergleich der *TDS-Tc*-Sequenz mit dem Codongebrauch von *S. cerevisiae* ergab Mangelcodons an 14 Stellen, die zu einer ineffizienten Translation führen könnten (Abbildung III.8). Daher wurde von Genescript das für die Translation in *S. cerevisiae* optimierte TDS-Gen *TDSco* synthetisiert. Dieses zeigte einen von 0,067 auf 0,59 erhöhten *Codon Adaptiveness Index* CAI<sup>[244]</sup> und führte bei Expression in CEN 7 nicht nur zu einer detektierbaren Proteinakkumulation (Abbildung III.6), sondern auch zu einer um das 40 fache gesteigerten Taxadienproduktion (Tabelle III.4). Dieser Wert stellt die bis heute höchste bekannte rekombinante Taxadienausbeute dar (Tabelle IV.2).

Tabelle IV.2: Vergleich der rekombinanten Taxadienproduktion verschiedener Expressionssysteme

Expressionssystem	Überexprimierte Gene	Taxadien [mg/l]	Referenz
<i>A. thaliana</i>	<i>TDS</i>	0,6	[147]
<i>E. coli</i>	<i>TDS</i> ; <i>GGPPS</i> ; <i>dxs</i> , <i>idi</i>	1,3	[142]
<i>S. cerevisiae</i>	<i>TDS</i> ; <i>GGPPS</i>	1	[149]
<i>S. cerevisiae</i>	<i>TDSco</i> ; <i>GGPPS-Sa</i> ; <i>tHMGR</i> ; <i>UPC2.1</i>	8,7	diese Studie

**TDS:** pflanzliches Taxadiensynthasegen; **GGPPS:** pflanzliches Geranylgeranyldiphosphatsynthasegen; **dxs:** DXP-Synthasegen; **idi:** IPP-Isomerasegen; **TDSco:** für die Expression in *S. cerevisiae* optimiertes TDS; **GGPPS-Sa:** GGPP-Synthasegen aus *S. acidocaldarius*; **tHMGR1:** Gen für die um die regulatorische Domäne verkürzte HMG-CoA-Reduktase Isozym 1 aus *S. cerevisiae*; **UPC2.1:** punktmutierte Version des Transkriptionsfaktors UPC2 aus *S. cerevisiae*, ermöglicht die Aufnahme von Sterol auch unter aeroben Bedingungen

Verschiedene Ansätze zur weiteren Optimierung der Isoprenoidbiosynthese in *S. cerevisiae* wurden im Laufe dieser Arbeit veröffentlicht, die auch an dem hier generierten Taxadien produzierenden Stamm CEN 7 angewendet werden könnten. So integrierten Nielsen und Mitarbeiter für die rekombinante Synthese sesquiterpenoider Verbindungen zusätzlich den bakteriellen DXP-Weg.<sup>[236]</sup> Die Gruppe um Shiba modifizierte zur Steigerung der Amorphadienproduktion hingegen den hefeeigenen Zentralstoffwechsel für eine bessere Bereitstellung von Acetyl-CoA, der Ausgangssubstanz des Mevalonatwegs.<sup>[239]</sup> Auch Strategien zur Umleitung der synthetisierten Isoprenoide weg vom Ergosterolbiosyntheseweg durch Repression der Squalensynthase führten



bereits zur gesteigerten Produktion des gewünschten Terpens.<sup>[240][241][245][246]</sup> Allerdings konnte im Rahmen dieser Arbeit bei Expression von *GGPPS-Sa* nur eine mit dem Wildtyp vergleichbare Squalenakkumulation beobachtet werden, was gegen eine übermäßige Synthese unerwünschter Nebenprodukte spricht. Eine Umleitung des Kohlenstoffflusses durch Repression der Squalensynthese würde daher vermutlich zu einer Reduktion des Ergosterolspiegels bis unterhalb der für das Zellwachstum benötigten Mengen führen. Weiterhin könnten weitere Kopien von *GGPPS-Sa* integriert oder die katalytische Aktivität des Enzyms wie von Leonard und Mitarbeitern durch Mutagenese optimiert werden.<sup>[247]</sup> Allerdings liegt mit 33 mg/L an GGOH (Tabelle III.5) der Engpass nicht im Primärmetabolismus oder der Synthese von GGPP, sondern in der von der Taxadiensynthese katalysierten Reaktion. Bei partieller Aufreinigung aus einer *Taxus*-Zellkultur zeigte die Taxadiensynthese mit  $0,4 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ <sup>[83]</sup> eine spezifische Aktivität, die nur einem Zehntel derjenigen partiell aufgereinigter *GGPPS-Sa* entspricht ( $8,7 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ ,<sup>[210a]</sup>). Die detektierte GGOH-Akkumulation könnte daher in einer guten Umsetzung der Isoprenoidvorstufen zu GGPP bedingt sein, welches aufgrund der niedrigen spezifischen Aktivität von TDS nicht quantitativ für die Synthese von Taxadien verwendet wird.

Das synthetisierte GGPP ist nicht nur Substrat der Taxadiensynthese, sondern wird daneben auch durch hefeeigene Phosphatasen zu GGOH dephosphoryliert.<sup>[248]</sup> Bereits 1996 konnte die Phosphataseaktivität von *S. cerevisiae*-Zellextrakten auf  $4,5 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$  bestimmt werden.<sup>[249]</sup> Die spezifische Aktivität der Taxadiensynthese beträgt  $6,997 \text{ nmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ <sup>[91]</sup> und ist damit nur geringfügig höher. Allerdings liegt das Enzym in der Zelle durch die Überexpression von *TDSco* in vielfacher Konzentration vor, sodass eine Umsetzung von GGPP zu Taxadien trotz des niedrigen  $k_{\text{cat}}$  der Taxadiensynthese und der parallelen Dephosphorylierung von GGPP zu GGOH möglich ist.

Die hefeeigene Phosphatase DPP1 ist verantwortlich für 75 % der Dephosphorylierung von GGPP und FPP,<sup>[248]</sup> sodass eine Deletion von *DPPI* die Umwandlung zu GGOH verringern könnte. Ein entsprechender Deletionsstamm zeigte jedoch eine niedrigere Terpenakkumulation als der Vergleichsstamm mit intaktem *DPPI*,<sup>[245]</sup> weshalb dieser Ansatz für die weitere Optimierung der Taxadienproduktion wenig aussichtsreich wäre.

*S. acidocaldarius* kann sowohl FOH, als auch GGOH jeweils zu FPP und GGPP phosphorylieren.<sup>[250]</sup> In CEN 7 wird das konstitutiv gebildete GGPP zu GGOH dephosphoryliert und parallel zu Taxadien zyklisiert, sodass von einer stationären Konzentration von GGPP in der Hefezelle ausgegangen werden kann. Durch die Expression einer entsprechenden Kinase aus *S. acidocaldarius* könnte das dephosphorylierte GGPP wieder phosphoryliert und die stationäre



Konzentration von GGPP erhöht werden. Trotz ihrer niedrigen Wechselzahl wäre die Taxadiensynthese dadurch in der Lage, mehr GGPP zu Taxadien umzusetzen.

Dass die Zyklisierung von GGPP zu Taxadien der limitierende Schritt ist, könnte neben dem relativ niedrigen  $k_{\text{cat}}$  der TDS von 0,0106 außerdem entgegen der bekannten *in vitro*-Stabilität<sup>[91]</sup> in einem raschen Abbau oder einer generellen Instabilität der TDS *in vivo* begründet sein. Es gibt keine Kristallstruktur der Taxadiensynthase, die eine rationale Evolution zu einer aktiveren oder stabileren Taxadiensynthese erleichtern würde. Die Mutagenisierung von *TDSco* mit anschließender Selektion auf bessere Taxadienproduzenten stellt daher eine Möglichkeit zur Steigerung der Produktausbeute dar. Dennoch katalysieren Terpenzyklen eine der komplexesten natürlich vorkommenden Reaktionen,<sup>[251]</sup> sodass eine Verbesserung der spezifischen Aktivität schwierig erscheint. Des Weiteren konnte bereits an verschiedenen Terpenzyklen gezeigt werden, dass sich durch den Austausch einer einzigen Aminosäure die Produktspezifität verändert.<sup>[252][253]</sup> Die Gefahr ist daher gegeben, auch im Falle einer Mutagenisierung der Taxadiensynthase deren Produktspezifität zu verlieren, anstatt ihre spezifische Aktivität zu erhöhen.

## IV.2 Generierung eines Genclusters

Im Rahmen dieser Arbeit wurde zunächst erfolgreich ein Hefestamm generiert, der über die Fähigkeit verfügt, Taxadien *de novo* in einem Titer zu synthetisieren, der den bisher erreichten um eine Zehnerpotenz übertrifft.<sup>[149]</sup> Dafür wurde der Hefestamm CEN.PK2-1C mit vier Vektoren transformiert, die jeweils eins der Gene *TDSco*, *GGPPS-Sa*, *tHMGR-Sc* und *upc2.1* enthielten. Die Stabilität der Vektoren wurde durch Komplementation der Auxotrophien von CEN.PK2-1C für die Synthese von Uracil, Histidin, Tryptophan und Leucin gewährleistet. In Ermangelung zusätzlicher Auxotrophien des Hefestamms war seine Verwendung für die Integration weiterer Gene aus dem Taxolbiosyntheseweg jedoch ohne den Gebrauch dominanter Selektionsmarker nicht möglich. Um diese Limitierung zu umgehen, wurden *TDSco-80*, *GGPPS-Sa* und *tHMGR1* jeweils inklusive eines eigenen Promoters und Terminators fusioniert und die Gene über einen Vektor koexprimiert. Die *de novo*-Synthese von Taxadien durch die Expression der fusionierten Gene auf einem einzelnen Plasmid benötigt nur einen einzigen Selektionsmarker und gestattet die Propagierung von drei weiteren Vektoren in den gleichen Stamm. Diese können für die Expression zusätzlicher, ebenfalls fusionierter Gene aus dem Taxolbiosyntheseweg verwendet werden und ermöglichen damit die rekombinante Produktion entsprechend funktionalisierter Taxane.

### IV.2.1 Fusionierung von DNS-Fragmenten

Zur Fusionierung von DNS-Fragmenten wurden bereits diverse Methoden entwickelt (Tabelle IV.3). Sie lassen sich grundsätzlich in Ligation nach Restriktion mit entsprechenden Endonukleasen,<sup>[254]</sup> SOE-PCR (*Splicing by overlap extension-PCR*)<sup>[160][255][256]</sup> und *in vivo*- bzw. *in vitro*-Rekombinationsmethoden einteilen.

**Tabelle IV.3: Methoden zur Fusionierung von DNS-Fragmenten**

Methoden	Beschreibung	Referenz
Restriktion/Ligation	Schneiden von DNS-Fragmenten durch Restriktionsendonukleasen in jeweiliger Erkennungssequenz, Verknüpfung gleich verdauter Fragmente durch Ligaseaktivität	[254]
SOE-PCR	Verbindung von DNS-Fragmenten über homologe überlappende Bereiche durch Denaturierungs-/Anlagerungs-/Elongations-Zyklus einer PCR	[160] [255] [256]
<b><i>In vivo</i> Rekombination</b>		
Domino-Methode	Homologe Rekombination zwischen überlappenden Sequenzen in <i>Bacillus subtilis</i> durch RecA-abhängigen Reparaturmechanismus	[225, 257]
DNA-Assembler	Homologe Rekombination zwischen 40bp-Überhängen in <i>S. cerevisiae</i>	[258]
<b><i>In vitro</i> Rekombination</b>		
SLIC ( <i>Sequence and Ligation Independent Cloning</i> )	Homologe Rekombination zwischen 30-40bp Überhängen durch T5-Exonuklease-Aktivität, Anlagerung homologer Bereiche, Auffüllen durch Polymerase und Reparatur mittels T4-Ligase	[259] [226]
USER-Cloning ( <i>Uracil Specific Excision Reagent</i> )	Einbau von Deoxyuridinbase (dU) in homologen Bereich (10-15 bp), Excision von dU durch Enzymmix von NEB (Uracil DNS Glykosylase und Endonuklease), Anlagerung der ss-DNS-Bereiche, Transformation in <i>E. coli</i> für Auffüllen und Ligation	[260]
In-Fusion	Generieren von ss-DNS-Überhängen durch 3'-Exonuklease-Aktivität von Poxvirus-DNS-Polymerase (Clontech, Mountain View, CA, USA), Anlagerung der homologen ss-DNS-Bereiche, Transformation in <i>E. coli</i> für Auffüllen und Ligation	[261]
Stargate	Integration von DNS-Fragmenten in Vektor über IBA Erkennungssequenzen, danach gerichtete Fusionierung von Inserts über Linkerregion durch Rekombination der Vektoren	BioTAGhology

Anhand homologer Rekombinationstechnik *in vivo* wie der Dominomethode konnte bereits ein DNS-Abschnitt von insgesamt 134 kb fusioniert werden.<sup>[225]</sup> Die *DNA assembler*-Methode wurde dafür verwendet, ganze Biosynthesewege einer Länge zwischen 9 und 19 kb in einem Schritt zu generieren.<sup>[258]</sup> Mithilfe der Kombination von *in vitro*- und *in vivo*-Rekombination baute die Gruppe

um Gibson das komplette *Mycoplasma genitalium*-Genom einer Größe von ca. 600 kb aus chemisch synthetisierten Oligonukleotiden zusammen.<sup>[262]</sup> Im Rahmen dieser Arbeit erfolgte die Generierung des Genclusters mittels der Kombination von SOE-PCR, Stargate- und Gateway-Technologie. Sowohl die über Stargate als auch über Gateway realisierten Fusionierungsschritte zeigten bei Analyse durch Testrestriktionen eine Effizienz von 100 %. Die fehlerfreie Fusionierung wurde dabei anhand von Sequenzierungen der jeweiligen Endkonstrukte bestätigt. Die verwendete Methode stellt daher einen erfolgreichen Ansatz zur Verlinkung von DNS-Fragmenten dar.



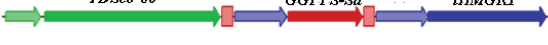






#### IV.2.2 Analyse generierter Fusionsstämme

Im Rahmen dieser Arbeit wurden neun Fusionskonstrukte der drei Gene *TDSco-80*, *GGPPS-Sa* (bzw. *GGPPSco*) und *tHMGR-Sc* inklusive jeweiliger Promoter- und Terminatorregionen generiert (Tabelle III.10). Sie variieren sowohl in der Orientierung der Gene zueinander, der Wahl der verwendeten Regulationseinheiten, sowie in der Sequenz des mittleren Gens *GGPPS*. Keines dieser Konstrukte führte zu einer gleichwertig detektierbaren Transkription aller drei Gene, obwohl dies bei Expression im jeweiligen Einzelvektor zweifelsfrei nachgewiesen werden konnte (Abbildung III.18).

Die starke Squalenproduktion mancher Fusionsstämme (CEN 8, CEN 9, CEN 12, CEN 13, CEN 15, Abbildung IV.2) kann wie auch bei Einzelexpression der Gene (IV.1) durch die Transkription von *tHMGR* bei gleichzeitig nur schwach oder gar nicht transkribiertem *GGPPS-Sa* erklärt werden. Da *tHMGR* ein dereguliertes Enzym des Mevalonatwegs darstellt, führt es zu einer verstärkten Synthese der Isoprenoidvorstufen IPP und DMAPP. Diese werden bei fehlender *GGPPS-Sa* nicht zu GGPP, sondern durch die hefeeigenen Enzyme GPP-, FPP- und Squalensynthase zu Squalen kondensiert. Die geringen Mengen an rekombinant produziertem GGPP durch die ineffiziente Transkription von *GGPPS-Sa* erklären weiterhin die niedrige Taxadienkonzentration, die wegen mangelnden Substrats trotz effizienter Transkription und Translation von *TDSco-80* ermittelt wurde (Abbildung IV.2).

Bei keinem der Fusionsstämme konnte *GGPPS*-mRNS nachgewiesen werden. Dies ist bei manchen Konstrukten mit dem Ansatz der transkriptionalen Interferenz zu vereinbaren.<sup>[213][219][220]</sup> Dabei würde bei einer Tandemorganisation der fusionierten Gene der jeweils *GGPPS* vorgeschaltete Promoter durch die Expression des 5'-gelegenen *TDSco-80* blockiert, was die Bindung der benötigten Transkriptionsfaktoren beeinträchtigt und zu der beobachteten geringen oder fehlenden

GGPP-Produktion führen könnte (CEN 8, CEN 10, CEN 14, CEN 16, Abbildung IV.3). Diese Theorie basiert auf der aktiven Expression von *TDSco-80*, die in jedem Fusionsstamm mittels Northernblot nachgewiesen werden konnte. Nicht nur *TDSco-80* und *GGPPS*, sondern auch *GGPPS* und *tHMGR1* sind dabei im Tandem orientiert. Die Blockierung der *GGPPS*-Expression würde allerdings eine Transkription des 3'-gelegenen *tHMGR1* zulassen (CEN 8).

		Taxadien [mg/l]	GGOH [mg/l]	Squalen [mg/L]
CEN 8		0,41±0,02	0,34±0,05	86,3±12
CEN 9		0,42±0,05	0,45±0,04	94,8±21
CEN 10		n.d.	0,25±0,01	3,20±0,1
CEN 11		0,78±0,01	0,23±0,03	3,12±0,2
CEN 12		1,86±0,1	2,5±0,1	26,5±7
CEN 13		0,41±0,01	4,9±0,1	47,1±2,7
CEN 14		n.d.	0,23±0,01	3,30±0,1
CEN 15		n.d.	0,64±0,01	119±24
CEN 16		0,45±0,01	0,36±0,01	4,00±0,1

**Abbildung IV.3: Übersicht generierter Fusionsstämme, der jeweiligen integrierten Gencluster und Produktakkumulation.** Promotorregionen sind als schraffierte Pfeile dargestellt (grün: gpdp; rot: gap1p; blau: adh1p; rosa: pgkp; hellblau: tefp; mangenta: pgkp aus pVV200). Terminatorregionen sind als schattierte Boxen angegeben (grün: pho5tt; rot: gcn4tt; blau: cyc1tt; mangenta: cyc1tt aus pVV200; rosa: adh1tt). Gelb: Peptidlinker.

Nicht vereinbar mit dieser Hypothese bleibt letztlich die Squalenproduktion der Fusionsstämme CEN 10, CEN 14 und CEN 16, die trotz inhibierter Transkription von *GGPPS* nur eine mit dem Wildtypstamm vergleichbare Konzentration erreicht und daher gegen eine Überexpression von *tHMGR1* spricht. Wie schon bei den Tandemkonstrukten, sind auch bei konvergent/divergent orientierten Fusionskonstrukten einige der Untersuchungsergebnisse mit auftretender

transkriptionaler Interferenz zu erklären. Die divergenten Expressionseinheiten *GGPPS* und *tHMGR1* könnten sich durch eine Promotorblockade einseitig (die Bindung der Polymerase an *adh1p* vor *tHMGR1* inhibiert diejenige an *gap1p* vor *GGPPS*, CEN 9) oder gegenseitig (CEN 11) reprimieren (Abbildung IV.3). Allerdings gibt es bei genauer Betrachtung der Analyseergebnisse einzelner Fusionsstämmе grundsätzliche Unstimmigkeiten: Die mittels GC/MS detektierten Produktmengen sind nicht immer mit den Resultaten durchgeführter Northern- und Westernblots vereinbar. So zeigten die Stämme CEN 13 und CEN 14 die gleiche Expression von *TDSco-80*, *GGPPS-Sa* und *tHMGR1*, aber eine stark unterschiedliche Produktion der durch die jeweiligen Enzyme katalysierten Synthese von Taxadien, GGOH und Squalen. Des Weiteren war in vielen Fällen trotz detektierbarem Transkriptionsprodukt von *tHMGR1* eine Translation auf Proteinebene nicht nachweisbar, die Aktivität von tHMGR1 wiederum anhand einer erhöhten Squalenproduktion indirekt bestätigt (CEN 8, CEN 9, CEN 12, CEN 13, CEN 15, Abbildung IV.3). Auch führt die Verwendung des gleichen Promoters in gleicher Position bei verschiedenen Konstrukten zu unterschiedlicher Produktmenge des entsprechenden Enzyms. So ist der *TDSco-80* vorgeschaltete Promoter in allen Konstrukten *gdpd* aus dem Vektorbackbone von pAG423. Bei jedem der generierten Fusionsstämmе konnte sowohl die Transkription von *TDSco-80*, als auch das entsprechende Translationsprodukt TDSco-80 nachgewiesen werden. Allerdings variierte die quantifizierte Taxadienmenge unabhängig von der detektierten GGOH-Menge. So synthetisierte CEN 11 fast 800 µg/L an Taxadien, obwohl sich die GGOH-Konzentration gegenüber dem Wildtypstamm nur leicht erhöhte. Die Produktanalyse von CEN 13 dagegen ergab trotz einer 12-fach höheren GGOH-Menge die gleiche Taxadienmenge wie in CEN 8 und CEN 9. Auch andere Promotoren hatten variierende Auswirkungen. So unterscheiden sich die Fusionskonstrukte der Stämme CEN 11 und CEN 12 ausschließlich in der Orientierung der divergenten Promotoren *pgkp-tefp* zwischen *GGPPS-Sa* und *tHMGR1*. Während die Konstellation *pgkp-tefp* verglichen mit dem Wildtyp weder zu einer gesteigerten Squalen-, noch GGOH-Produktion führte, resultierte die Promotororientierung *tefp-pgkp* in der höchsten Produktkonzentration sowohl an GGOH als auch an Squalen aller generierten Fusionsstämmе (Abbildung IV.3). Auch die Terminator-Promoterregion *gcn4tt-adh1p* erwies sich in Tandemkonstrukten bei Positionierung vor *tHMGR1* in manchen Konstrukten als funktionell (CEN 8, CEN 16), führte in CEN 10 allerdings zu keiner tHMGR-bedingten gesteigerten Isoprenoidvorstufensynthese und dementsprechend hohen Squalenakkumulation.

#### IV.2.3 *Instabilität der Fusionsvektoren in der Hefe*

Bei keinem der generierten Fusionsstämme sind die erhaltenen Resultate hinsichtlich analysierter Transkription, Translation und Produktbildung miteinander vereinbar. Allerdings konnte gezeigt werden, dass sich die jeweiligen Fusionskonstrukte nach Transformation in die Hefezelle verändern (Abbildung III.32). Dafür wurde aus jedem Fusionsstamm Gesamtplasmid isoliert, in *E. coli* vereinzelt und durch die Restriktionsendonuklease *XbaI* verdaut. Das Fragmentierungsmuster der gelelektrophoretisch aufgetrennten Testrestriktionen entsprach nur in einem Fall (CEN 11) dem ursprünglich transformierten (Abbildung III.32). Anhand einer Zeitreihe, im Rahmen derer nach 24, 48 und 72 h jeweils Gesamtplasmid isoliert und vereinzelt wurde, konnte eine von der Inkubationszeit unabhängige Veränderung demonstriert werden, die weiterhin bei allen untersuchten Fusionsvarianten auftrat (Abbildung III.33). Der Verdau von jeweils drei Vektoren der isolierten Plasmidfraktion einer Kultur führte dabei zu jeweils völlig unterschiedlichen Fragmentierungsmustern, was darauf hindeutet, dass die Veränderung der Fusionskonstrukte innerhalb der Hefezelle nicht einheitlich erfolgt (Abbildung III.33). Gestützt wird diese Theorie durch die Sequenzierung eines der reextrahierten Plasmide. Der Vergleich der erhaltenen Basenabfolge mit derjenigen des ursprünglich in die Transformation eingesetzten Fusionskonstrukts zeigte das Fehlen des gesamten Genclusters (Abbildung III.34). Eine einheitliche Veränderung der Plasmide innerhalb einer Hefekultur würde in diesem Fall im Verlust der Gensequenzen für *TDSco-80*, *GGPPS-Sa* und *tHMGR* aller Plasmide resultieren. Die Untersuchung des Stammes ergab allerdings sowohl Transkriptionsprodukt von *TDSco-80* und *tHMGR1* (Abbildung III.15), eine detektierbare Akkumulation von *TDSco-80* (Abbildung III.14), als auch eine im Vergleich zum Wildtyp erhöhte Produktion von Taxadien, GGOH und Squalen (Tabelle III.6). Neben dem fehlenden Gencluster zeigte die Sequenzanalyse des reextrahierten Plasmids zusätzlich den Verlust des Histidinmarkers. Das Fusionskonstrukt war aus einer Hefekultur isoliert worden, deren Inkubation in Selektionsmedium erfolgte, sodass ein Wachstum ausschließlich durch die gleichzeitige Propagation eines Plasmids mit Selektionsmarker ermöglicht werden konnte.

Die Veränderung der Fusionsplasmide in der Hefezelle erfolgte bereits innerhalb der ersten 24 Stunden (Abbildung III.33). Die Analyse der generierten Stämme auf Transkriptions-, Translations- und Produktebene wurde dagegen anhand von RNS-, Protein- und Zellextrakten in Einzelexperimenten durchgeführt, die nach einem längeren Inkubationszeitraum aus der jeweiligen Hefekultur isoliert wurde (II.1.3.2; II.2.1.3; II.3.1.1). Daher kann auch bei diesen Ergebnissen von keiner einheitlichen Plasmidfraktion in der Zelle ausgegangen werden. Die erhaltenen Northern- und Westernblots, bzw. GC/MS-Daten basieren dementsprechend auf Durchschnittswerten



heterogener Hefezellen, die zu unterschiedlichen Zeitpunkten untersucht wurden und somit weder auf Transkriptions- oder Translationsebene, noch in ihrer Produktakkumulation übereinstimmen müssen. Eine schlüssige Interpretation erhaltener Daten der jeweiligen Fusionsstämme ist daher unmöglich.

#### IV.2.3.1 Mögliche Ursachen der Plasmidveränderung

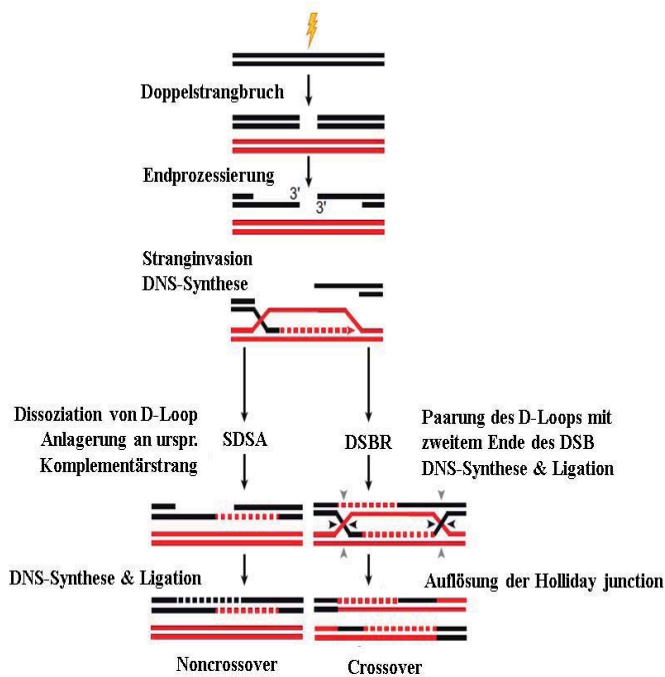
Die Veränderung von DNS-Sequenzen *in vivo* kann anhand verschiedener Rekombinationsmechanismen, die durch Strangbrüche der DNS initiiert werden, erfolgen. Solche Strangbrüche können sequenzunabhängig durch einen Zusammenbruch der Replikationsgabel, freie Radikale, Chemikalien oder Strahlung ausgelöst werden.<sup>[263]</sup> Weiterhin sind mehrere DNS-Strangbruch-katalysierende Enzyme in *S. cerevisiae* wie die HO-Endonuklease beim Wechsel des Paarungstypen,<sup>[264]</sup> Homing-Endonukleasen wie PI-SceI<sup>[265]</sup> oder I-SceI und II,<sup>[266]</sup> die Flp-Rekombinase bei Amplifikation des 2 µm-Plasmids,<sup>[189]</sup> Rad1-Rad10 und Rad2 als Endonukleasen beim Austausch falsch gepaarter Nukleotide,<sup>[267]</sup> während der Meiose, sowie Topoisomerasen während der DNS-Replikation<sup>[268]</sup> bekannt. Für die Reparatur spontaner und katalysierter DSB existieren zahlreiche Mechanismen, die sich grundsätzlich in homologe und nicht-homologe Rekombination einteilen lassen.

Die homologe Rekombination gilt als universeller Reparaturmechanismus für den Erhalt der genomischen Integrität nach einem Doppelstrangbruch und verkörpert einen präzisen Reparaturmechanismus, der auf homologen DNS-Sequenzen basiert. Obwohl homologe Rekombination (HR) in allen untersuchten Organismen nachgewiesen werden konnte,<sup>[269]</sup> wurden die Studien zu ihrem Mechanismus hauptsächlich an Hefen durchgeführt.<sup>[270]</sup> HR kann dabei zwischen Schwesterchromatiden, Allelen homologer Chromosomen, nicht allelen homologen Sequenzen wie Plasmiden oder exogenen, linearen DNS-Fragmenten, sowie verstreuten Kopien in verschiedener genomischer Lage auftreten.

Der grundsätzlich anerkannte Mechanismus der homologen Rekombination ist eine Kombination zweier Modelle für die Reparatur von Doppelstrangbrüchen<sup>[271]</sup> basierend auf den beobachteten Produkten nach vollendeter Reparatur *Crossing over* (CO) und Genkonversion (*Non crossing over* NCO, Abbildung IV.4). Das erste Modell wurde von Szostak und Mitarbeitern unter dem Namen **Doppelstrangbruchreparatur** (DSBR) veröffentlicht.<sup>[272]</sup> Dabei wird die HR durch einen DSB initiiert und die freien Enden durch Nukleaseaktivität so prozessiert, dass einzelsträngige



3'-Überhänge entstehen. Nach erfolgreicher Homologiesuche lagert sich dieser Überhang an die komplementäre Sequenz an. Diese als Stranginvasion bezeichnete Verdrängung des ursprünglichen Komplementärstrangs führt zu der Bildung eines D-Loops (*displacement Loop*). Durch die Verlängerung des 3'-Endes mittels DNS-Polymeraseaktivität vergrößert sich das D-Loop-Intermediat und paart sich mit dem 5'-Ende des Doppelstrangbruchs. Die dabei gebildete doppelte *Holliday Junction* (HJ) wird durch Endonukleasen aufgelöst und führt zur Trennung der DNS-Stränge und einem Crossover der beteiligten DNS-Abschnitte.



Quelle: modifiziert nach <sup>[273]</sup>

**Abbildung IV.4: Schematische Darstellung zweier Modelle homologer Rekombination <sup>[271]</sup>:**

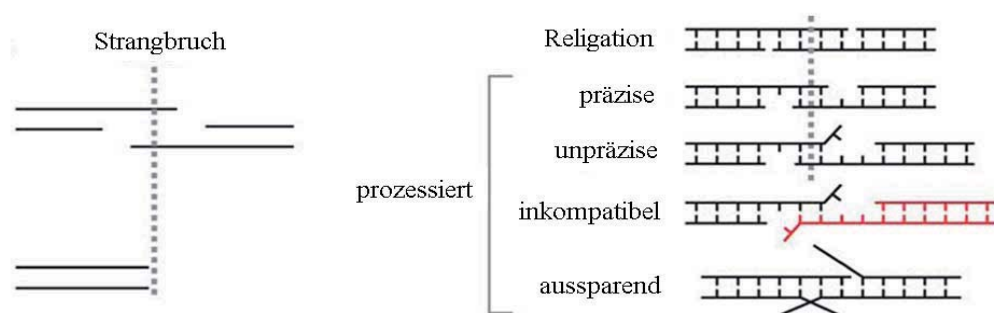
Homologe Rekombination wird durch einen Doppelstrangbruch initiiert. Durch Nukleaseaktivität werden die Enden prozessiert, wodurch einzelsträngige 3'-Überhänge entstehen. Nach abgeschlossener Homologiesuche paart sich dieser Überhang mit der komplementären Sequenz und wird durch DNS-Polymeraseaktivität verlängert. Im Falle des **DSBR (rechts)** bindet der entstandene Displacement Loop (D-Loop) an das 5'-Ende des Doppelstrangbruchs und bildet eine doppelte Holliday junction, die durch asymmetrische Trennung aufgelöst wird. Das Produkt ist ein Crossover der beiden ursprünglichen DNS-Stränge. Bei **SDSA (links)** löst sich der nach Invasion verlängerte Strang wieder und paart sich mit dem anderen Ende des Doppelstrangbruchs ohne die Bildung einer Holliday junction und anschließendem Crossover.

Alternativ können sich die beiden Enden des Doppelstrangbruchs unabhängig voneinander an homologe Sequenzen anlagern. Nach Verlängerung werden sie wieder freigesetzt und paaren sich miteinander ohne die Formation einer intermediären Holliday Junction mit anschließendem Crossover (Abbildung IV.4 links). Dieses Modell wird als **syntheseabhängige Stranganlagerung** (*Synthesis-Dependent Strand annealing* SDSA) bezeichnet.<sup>[274]</sup> Es wird angenommen, dass HR, die in Genkonversion resultiert, ausschließlich über SDSA erfolgt, während Crossoverprodukte die einzigen Produkte der DSBR darstellen (Abbildung IV.4).

Obwohl homologe Rekombination stets auf einer Paarung gleicher Sequenzen beruht, ist das Ergebnis dieses Reparaturmechanismus nicht immer homolog zur ursprünglichen Basenabfolge des betroffenen DNS-Abschnitts. Tritt ein Doppelstrangbruch zwischen zwei repetitiven Sequenzen

(*direct repeats*) von mehr als 30 Basenpaaren auf, können die prozessierten Enden über die komplementären *repeats* aneinanderbinden. Die 3'-Überhänge werden nukleolytisch entfernt und die entstandenen Enden ligiert. Die als *Single-Strand Annealing* (SSA) bezeichnete Reparatur führt zu einer Deletion des DNS-Abschnitts zwischen den repetitiven Sequenzen und ist unabhängig von einer Stranginvasion oder Bildung einer Holliday Junction.<sup>[275]</sup>

Im Gegensatz zur homologen Rekombination basiert nicht-homologe, bzw. illegitime Rekombination auf einer Reparatur von Strangbrüchen ohne die Beteiligung homologer Sequenzen (*Non-Homologous End Joining* NHEJ)<sup>[276]</sup> oder auf einer Verbindung über sogenannte mikrohomologe Bereiche von 5-25 Nukleotiden (*Microhomology Mediated End Joining* MMEJ).<sup>[277]</sup> Verschiedene Ursachen eines Strangbruchs führen dabei zu unterschiedlichen Konfigurationen des Bruchs. Eine unterbrochene DNS-Replikation oder Chromosomenbrüche in Telomerbereichen resultieren in Strangbrüchen mit nur einem freien Ende (Abbildung IV.5 links oben). Die Folge einer Beschädigung der DNS durch chemische Agenzien oder Strahlung hingegen sind DSB mit gepaarten Enden. (Abbildung IV.5 links unten). Mit Ausnahme der Religation unprozessierter Enden oder präziser Aneinanderlagerung entstandener Überhänge führt die nicht-homologe Rekombination zu Deletionen von Bereichen unterschiedlicher Länge (Abbildung IV.5 rechts).



**Abbildung IV.5: Schematische Darstellung nicht-homologer Rekombinationstypen<sup>[276]</sup>.** Links: Konfigurationstypen von Strangbrüchen. Rechts: mögliche Reparaturprodukte

Anhand nicht-homologer Reparaturmechanismen können nicht nur die bereits beschriebenen Deletionen auftreten. Eine durch DNS-Strangbruch linearisierte Plasmidsequenz kann ebenso mit einer anderen verknüpft werden, was anstatt in einer Deletion in einer Insertion resultiert.

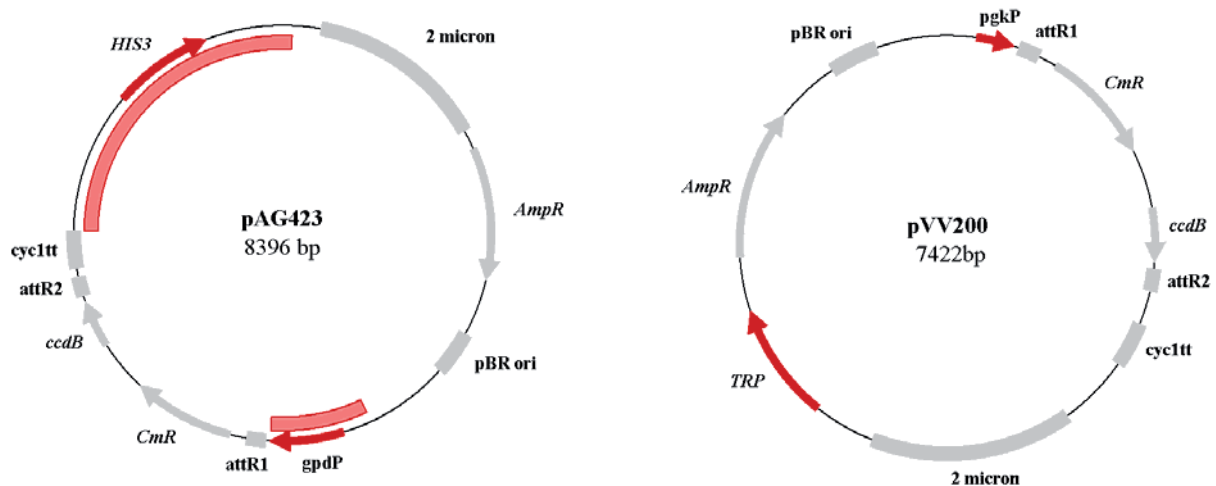
Die im Rahmen dieser Arbeit beobachtete Vielfältigkeit der Veränderung untersuchter Fusionskonstrukte spiegelt die verschiedenen Folgen möglicher nicht-homologer

Rekombinationsmechanismen wieder. So zeigte der Testverdau reextrahierter Vektoren (Abbildung III.33) Fragmentierungsmuster, die auf eine geringere Plasmidgröße schließen lassen als die ursprüngliche und somit Deletionen aufweisen (B1-11, B13-18, C1-8, C10-13, C15-18; D2, D4-8, D12-17). Ebenso entstanden Fragmentierungsmuster, die auf einen Fusionsvektor hindeuten, der größer ist als der in die Transformation eingesetzte, also inserierte Bereiche aufweist (C14, C18, D1, D10, D11). Auch reextrahierte Vektoren, die zwar vom ursprünglichen Muster abweichen, aber in der Gesamtgröße gleich bleiben, konnten beobachtet werden (D3, D9, D18). Eine mögliche Erklärung hierfür bietet die Veränderung der Basenabfolge innerhalb der Erkennungssequenz ohne Insertion oder Deletion.

Da die beobachtete Veränderung der Fusionsplasmide quantitativ verlief (Abbildung III.33), kann Rekombination aufgrund spontaner DNS-Strangbrüche ausgeschlossen werden. Die auftretende Häufigkeit lässt viel eher die Vermutung zu, dass es sich um eine enzymkatalysierte Reaktion handeln muss. Im Hinblick auf die Stabilität der Konstrukte in *E. coli* kann weiterhin angenommen werden, dass diese durch die Aktivität hefeeigener Enzyme modifiziert wurden. Die genetische Stabilität von Zellen ist abhängig von einem funktionierenden, auf verschiedenen Ebenen regulierten Rekombinationsapparat. Eine Mutation von Genen, die für involvierte Proteine kodieren, kann daher nicht nur zu einer fehlerhaften Reparatur auftretender DNS-Beschädigungen führen, sondern ebenfalls die Rekombinationsfrequenz verändern. Die beobachtete quantitative Plasmidveränderung der pAG-Vektoren könnte daher durch Veränderung oder Inaktivierung essenzieller Gene des Regulationsapparates und damit einhergehender Hyperrekombination entstanden sein. Anhand der Reextraktion und Fragmentierung durch Restriktionsendonukleasen konnte allerdings gezeigt werden, dass mit Ausnahme des Expressionsvektors pAG423 alle anderen verwendeten Plasmide (pRS- und pVV-Serie) unverändert in CEN.PK2-1C propagiert werden (Abbildung III.36). Eine vom Expressionsstamm abhängige Veränderung der Plasmid-DNS kann somit ausgeschlossen werden.

Ebenso kann eine Insert-abhängige Veränderung des Fusionsvektors ausgeschlossen werden. Sowohl die Integration eines einzelnen Gens in den Vektorbackbone von pAG423, als auch der Leervektor pAG423 führten nach Transformation zu einem vom ursprünglich in die Hefezelle transformierten Vektor abweichenden Restriktionsmuster (Abbildung III.35). Zusätzlich konnte am Beispiel von zwei Clustern demonstriert werden, dass diese im Gegensatz zu pAG423 bei Klonierung in den pVV200 auch nach Transformation in die Hefe wie der jeweilige Ursprungsvektor fragmentiert wurden. Die beobachtete Instabilität der generierten pAG-Fusionsplasmide basiert daher weder auf der Größe noch der Sequenz des integrierten Clusters.

Der einzige Unterschied zwischen dem stabilen CEN 17 und dem instabilen CEN 16, bzw. CEN 18 und CEN 12 liegt in den verwendeten Vektorgrundgerüsten pAG423 und pVV200. Diese variieren sowohl im Selektionsmarker für die Hefe (pAG423: Histidin, pVV200: Tryptophan), als auch im Promotor (pAG423: *GPD*-Promotor, pVV200: *PGK*-Promotor) und der Anordnung einzelner Elemente zueinander (pAG423: Selektionsmarker→2micron→AmpR; pVV200: 2micron→Selektionsmarker→AmpR). Abbildung IV.6 gibt die Vektorkarten der beiden Leervektoren pAG423 (Abbildung IV.6 links) und pVV200 (Abbildung IV.6 rechts) wieder. Rot gekennzeichnet sind die variierenden Sequenzabschnitte. Hier wird anschaulich, dass pAG423 sich nicht nur in dem verwendeten Hefeselektionsmarker selbst, sondern auch im Bereich um diesen herum (926 bp vor *HIS3* bis einschließlich 570 bp nach dem Selektionsmarker) von pVV200 unterscheidet. Diese Sequenz ist allerdings ebenso im pRS313 enthalten, der für die Generierung der Taxadien produzierenden Stämme CEN 4-7 verwendet wurde und nachweislich stabil ist (Abbildung III.36). Als Grund für die Instabilität von pAG423 kann die entsprechende Basenabfolge daher ausgeschlossen werden. Auch der zu pVV200 nicht homologe Sequenzabschnitt des *GPD*-Promotors in pAG423 wird in anderen Hefeexpressionsvektoren erfolgreich verwendet (pRS433-pRS436, pRS444-446, p426-GPD) und kann daher die Instabilität von pAG423 nicht erklären.



**Abbildung IV.6:** Vektorkarte des verwendeten Vektors pAG423. *AmpR*: Ampicillin-Resistenzgen; *gpdP*: Promoterregion des Glycerinaldehyd-3-phosphatdehydrogenasegens; *attR1*: Rekombinationssequenz für Gatewayklonierung; *CmR*: Chloramphenicol-Resistenzgen; *ccdB*: Gen für Gyrasehemmer; *attR2*: Rekombinationssequenz für Gatewayklonierung; *cyc1tt*: Terminatorregion des Cytochrom C1-gens; *HIS3*: Histidinmarker für die Selektion in Hefezellen; *pgkP*: Promoterregion des Phosphoglyceratkinasegens; *TRP*: Tryptophanmarker für die Selektion in Hefezellen; **grau**: Elemente, die sowohl im pAG423, als auch im pVV200 integriert sind; **rot**: Sequenzbereiche, in denen sich pVV200 und pAG423 unterscheiden (Pfeile: Promotorregion und Selektionsmarker in Hefen; Box: Gesamtbereich von pAG423, der nicht homolog ist zu pVV200)

Obwohl die einzelnen Elemente von pAG423 nicht für seine quantitative Veränderung in der Hefezelle verantwortlich zu sein scheinen, entstehen durch ihre Umordnung im Vergleich zu pVV200 neue Übergangsbereiche zwischen ihnen. Diese könnten Erkennungssequenzen DBS-katalysierender Enzyme enthalten, wodurch die Frequenz induzierter Rekombination signifikant erhöht würde. Eine Analyse dieser Bereiche ergab keine Basenabfolge, die den bekannten Schnittstellen hefeeigener Endonukleasen entspricht. Allerdings weist pAG423 mehrere DNS-Abschnitte einer Länge von 10-14 bp mit einer (Mikro-) Homologie zum restlichen Vektorgrundgerüst auf, die in pVV200 nicht vorhanden sind. Eine quantitative Rekombination zwischen diesen Bereichen basiert dennoch auf einer Initiierung durch Doppelstrangbrüche, die nicht spontaner Natur sein kann. Die beobachtete Frequenz der Plasmidveränderungen würde daher zusätzliche Faktoren benötigen, die die Plasmid-DNS an den entsprechenden Stellen quantitativ beschädigen.

Mehrere Fusionsplasmide wurden mithilfe PCR-amplifizierten Vektorgrundgerüsts generiert. Der Nachweis einer korrekten Basenabfolge reduzierte sich dabei ausschließlich auf das integrierte Gencluster. Mutationen innerhalb des nicht überprüften Bereichs des Vektors könnten daher zu einer erhöhten Rekombinationsfrequenz geführt haben. Obwohl diese Möglichkeit nicht auszuschließen ist, zeigte sich pAG423 auch als Leervektor in der Hefe instabil. Dieser war ausschließlich in *E. coli* vervielfältigt worden. Mit einer Mutationsrate in lebenden Organismen von  $10^{-8}$  bis  $10^{-11}$  pro Basenpaar<sup>[278]</sup> ist die Wahrscheinlichkeit einer quantitativen Veränderung des eingesetzten Vektors in *E. coli* allerdings ausgeschlossen.

Durch den Vergleich mit entsprechenden pVV200-Derivaten, die im gleichen Hefestamm stabil propagiert werden konnten, und der beobachteten Instabilität des pAG423-Leervektors in der Hefezelle, kann eine auf das Vektorgrundgerüst von pAG423 zurückzuführende Veränderung angenommen werden. Aufgrund der unterschiedlichen Restriktionsmuster von Fusionskonstrukten aus derselben Hefekultur kann weiterhin spekuliert werden, dass der Ort der Veränderung innerhalb des Vektors variiert und diese daher an keinem spezifischen DNS-Abschnitt erfolgt. Eine Sequenzierung weiterer reextrahierter Plasmide könnte darüber Aufschluss geben, ob die auftretende Plasmidveränderung dennoch einem bestimmten Muster folgt. Der Austausch des für die Expression des Genclusters verwendeten Vektors resultierte allerdings in stabilen Fusionsvektoren, sodass die Aufklärung der beobachteten Rekombinationsereignisse im Rahmen dieser Arbeit nicht zielführend war.

#### IV.2.4 Umklonierung des Clusters in pVV200

Alle für die Generierung des Taxadien produzierenden Stamms CEN 7 verwendeten Vektoren (Tabelle III.1) zeigten bei einem Verdau nach Reextraktion aus der Hefe das erwartete Fragmentierungsmuster (Abbildung III.36) und waren somit auch in der Hefe nachweislich stabil. Obgleich sowohl pVV200, als auch pVV214 2µm-basierte Vektoren darstellen und daher in der Hefe in gleicher Kopienzahl vorliegen, entsprach der Anteil von pVV200::*GGPPS-Sa* 87% aller reextrahierten Plasmide. Diese Menge spiegelt allerdings nicht die Anzahl der jeweiligen Vektoren in der Hefezelle wieder, sondern diejenige der Plasmidisolierung aus *E. coli*. Die Transformationseffizienz von Plasmiden in *E. coli* ist bekanntermaßen größenabhängig. So konnte bereits gezeigt werden, dass ein Plasmid von 2,7 kb zehnmal effizienter transformiert werden kann, als eins der 1,5-fachen Größe von 4,4 kb. Übersteigt die Plasmidgröße 20 kb, sinkt die Effizienz weiter auf ein Tausendstel.<sup>[279]</sup> Während pVV200::*GGPPS-Sa* mit *GGPPS-Sa* ein Gen der Länge von 1 kb exprimiert, ist in pVV214::*TDSco* das Taxadiensynthasegen integriert, welches im Vergleich zu *GGPPS-Sa* eine 2,5-fache Länge aufweist. Die beobachtete unterschiedlich effiziente Transformation der beiden Vektoren ist daher mit den Literaturwerten vereinbar.

Für eine stabile Expression in Hefen wurde sowohl das Fusionskonstrukt *TDSco-80-pho5tt-gcn4tt-GGPPS-Sa-tefp-pgkp-tHMGR1*, als auch *TDSco-80-pho5tt-tefp-GGPPSco-adh1tt-adh1p-tHMGR1* in den Destinationsvektor pVV200 rekombiniert und durch Transformation von CEN.PK2-1C die Stämme CEN 18 und CEN 17 generiert (Tabelle III.13). Im Gegensatz zu CEN 12 und CEN 16, in denen die jeweiligen Cluster über pAG423 exprimiert wurden, zeigten die aus CEN 17 und CEN 18 reextrahierten Fusionskonstrukte bei einem Testverdau das erwartete Fragmentierungsmuster (Abbildung III.37). Ebenso konnte bei Analyse der Proteinfractionen von CEN 18 und CEN 17 eine Akkumulation sowohl von *TDSco*, *GGPP-Sa* (bzw. *GGPPco*) und *tHMGR1* detektiert und somit die erfolgreiche Expression aller rekombinanten Gene nachgewiesen werden (Abbildung III.38). Die GC/MS-basierte Untersuchung von CEN 18 und CEN 17 auf ihre Produktion an Taxadien und GGOH ergab mit  $2,7 \pm 0,1$  mg/L und  $3 \pm 0,1$  mg/L, bzw  $15,4 \pm 0,3$  mg/L und  $12,4 \pm 0,3$  mg/L vergleichbare Mengen. Die Squalenkonzentrationen hingegen unterschieden sich mit  $31,3 \pm 3,2$  mg/L (CEN 18) und  $60,3 \pm 5,2$  mg/L (CEN 17) um den Faktor 2.

Während die Mengen detektierten Taxadiens und GGOHs auf eine vergleichbare Aktivität der entsprechenden Enzyme in den Stämmen hindeuten, lässt die stark abweichende Squalenproduktion auf eine unterschiedliche *tHMGR1*-Aktivität der beiden Stämme schließen. Die beiden Fusionsstämme unterscheiden sich in dem *tHMGR1* vorgeschalteten Promotor, sodass die



detektierten Squalenmengen in einer verschieden starken Expression des Gens begründet sein könnten. CEN 17 wurde mit einem Fusionsvektor transformiert, bei dem *tHMGR1* über den *PGK*-Promotor exprimiert wird. Für das Cluster in CEN 18 wurde dagegen der *ADHI*-Promotor für die Expression von *tHMGR1* gewählt. Obwohl beide Promotoren eine konstitutive Expression gewährleisten, führten sie auch in anderen Arbeitsgruppen zu unterschiedlichen Ergebnissen.<sup>[280]</sup> So wurde bei einem direkten Vergleich von *adh1p* und *pgkp* eine stärkere Aktivität des entsprechenden Genprodukts beobachtet, wenn das Gen durch *pgkp* exprimiert wurde. Obwohl Polakowski eine von der Kopienzahl exprimierter *tHMGR1* unabhängige tHMGR1-Aktivität beobachtete,<sup>[242]</sup> wurden auch abweichende Resultate veröffentlicht. So konnte Ro die Amorphadienproduktion durch Integration einer zweiten Kopie von *tHMGR1* um 50% steigern.<sup>[240]</sup> Ebenso war der Cubeboltiter bei Expression von *tHMGR1* über ein *high-copy* Plasmid höher als bei genomischer Integration des Gens.<sup>[241]</sup> Es ist daher nicht ausgeschlossen, dass auch im Rahmen dieser Arbeit eine stärkere Expression von *tHMGR1* in CEN 17 zu mehr Protein führt, was wiederum in einer effizienteren Synthese der Vorstufen IPP und DMAPP resultiert und diese schließlich zu einer größeren Menge von Squalen kondensiert werden als in CEN 18. Die Akkumulation von Squalen aufgrund einer Überexpression von *tHMGR1* trotz gleichzeitiger Expression von Terpenzyklasen wurde bereits mehrfach gezeigt.<sup>[69][242][241]</sup> Bedingt wird dies durch die kinetischen Eigenschaften der Squalensynthase (ERG9). Obwohl deren  $K_M$ -Wert von  $2,5 \mu\text{M}$ <sup>[281]</sup> mit dem bei Sesquiterpensynthasen beobachteten von  $0,4\text{--}10 \mu\text{M}$ <sup>[282]</sup> vergleichbar ist, unterscheidet sich ERG9 deutlich in seiner Wechselzahl. Der  $k_{\text{cat}}$  von  $0,53/\text{sec}$ <sup>[281]</sup> ist um zwei Zehnerpotenzen höher als bei der Amorpha-4,11-diensynthase, welcher je nach pH-Wert und Metallion zwischen  $3,4 \cdot 10^{-4}$  und  $1,54 \cdot 10^{-2}$  variiert.<sup>[282]</sup> Das synthetisierte FPP wird daher eher zu Squalen kondensiert, als für die Synthese von Sesquiterpenen verwendet. Obwohl das in dieser Arbeit verwendete bifunktionale Enzym GGPPS-Sa aus *Sulfolobus acidocaldarius* die Kondensation von IPP und DMAPP über GPP und FPP bis zum Endprodukt GGPP katalysiert, wurde auch in *S. acidocaldarius* eine Bereitstellung von FPP für die Synthese von Squalen nachgewiesen.<sup>[210a]</sup> Die beobachtete Menge an produziertem Squalen ist daher trotz Verwendung der bifunktionalen GGPP-Sa erklärbar. Eine Akkumulation von Squalen basiert allerdings nicht nur auf seiner aktiven Synthese, sondern ebenso auf einer weniger aktiven weiteren Umsetzung. Diese wird durch die Squalenepoxidase katalysiert. Während die Squalensynthase einen  $K_M$ -Wert von  $2,5 \mu\text{M}$ <sup>[281]</sup> und eine spezifische Aktivität von  $460 \text{ pmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ <sup>[283]</sup> aufweist, liegen die entsprechenden Werte für die Squalenepoxidase bei  $13,5 \mu\text{M}$ , bzw.  $32,1 \text{ pmol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ .<sup>[284]</sup> Es wird daher davon ausgegangen, dass die von der Epoxidase katalysierte Reaktion den regulierenden Schritt der Ergosterolbiosynthese darstellt.<sup>[285]</sup>



Die Fusionsstämme zeigen im Vergleich zu CEN 7, der die entsprechenden Gene einzeln exprimiert, eine 10- bzw. 20-fache Menge synthetisierten Squalens (Tabelle III.13). Neben der Verwendung unterschiedlicher Promotoren könnte dies durch die höhere Kopienzahl von *tHMGR1* in CEN 17 und CEN 18 bedingt sein. Während *tHMGR1* bei Einzelexpression in dem *low-copy*-Plasmid pRS315 integriert ist, wurden die jeweiligen Fusionskonstrukte in den *high-copy*-Vektor pVV200 rekombiniert, der in einem Vielfachen von pRS315 vorliegt.

Die verglichen mit CEN 7 signifikant niedrigere Produktion von GGOH der Fusionsstämme CEN 17 und 18 (50 bzw. 36%) könnte ebenfalls durch die unterschiedlichen Promotorregionen begründet sein. Während *GGPPS-Sa* bei Einzelexpression im pVV200 (CEN 7) der *PGK*-Promotor vorgeschaltet ist, wird sowohl *GGPPS-Sa* (CEN 18) als auch *GGPPSco* (CEN 17) im Cluster über den *TEF*-Promotor exprimiert, was zu einer schwächeren Expression führen könnte. Sowohl für die Analyse der *tHMGR1*- als auch der *GGPPS*-vorgeschalteten Promotoren, bzw. der unterschiedlichen verwendeten Plasmide könnte anhand einer RT-PCR die jeweilig synthetisierte mRNA quantifiziert und damit die Expressionsstärke der verschiedenen Promotoren verglichen, bzw. die Auswirkungen variierender Kopienzahlen bestimmt werden.

Während sich die Expression von *GGPPS* und *tHMGR1* in den Fusionsstämmen von derjenigen in CEN 7 in Kopienzahl des verwendeten Vektors oder der Wahl des vorgeschalteten Promotors unterscheidet, stimmen bezüglich der Expression von *TDSco-80* alle Stämme überein. Der jeweils verwendete Vektor ist sowohl bei Einzelexpression als auch im Falle des Clusters ein 2µm-Plasmid der pVV-Serie. Da *TDSco-80* innerhalb der Fusionskonstrukte das jeweils 5'-positionierte Gen darstellt, ist der ihm vorgeschaltete Promoter stets der vektoreigene *PGK*-Promotor. Die niedrige Menge synthetisierten Taxadiens der Fusionsstämme (2,7-3 mg/L) im Vergleich zu CEN 7 (8,7 mg/L) ist daher nicht durch unterschiedliche Promotorstärken oder eine abweichende Kopienzahl zu erklären. Ob die abweichende Taxadienproduktion dennoch auf einem variierenden Expressionslevel basiert, könnte wie auch bei Analyse der GGOH- und Squalenmengen anhand einer RT-PCR untersucht werden.

#### **IV.2.5 Expression des Clusters über ein künstliches Hefechromosom**

Die Integration des Genclusters zur Taxadienproduktion in ein modifiziertes, zirkuläres YAC führte nach Transformation von CEN.PK2-1C zu einem Hefestamm (CEN 19), der alle Gene des Konstrukts exprimiert (Abbildung III.40). Die Aktivität der entsprechenden Enzyme konnte

indirekt anhand des produzierten Taxadiens, GGOH und Squalens mittels GC/MS nachgewiesen werden (Tabelle III.14). Im Vergleich zu CEN 17, bei dem der identische Gencluster mithilfe des *high-copy*-Vektors pVV200 exprimiert wurde, konnte in CEN 19 nur ein Fünftel der Taxadien- bzw. ein Viertel der GGOH-Konzentration detektiert werden. Das Nebenprodukt Squalen hingegen wurde bei Expression des Fusionskonstrukts mithilfe des pYAC in einer fast doppelt so hohen Konzentration nachgewiesen (Tabelle III.14).

Eine Analyse der jeweiligen Relationen von Produkt zu Intermediat (Taxadien-GGOH) und Intermediat zu Nebenprodukt (GGOH-Squalen) der generierten Stämme CEN 7, CEN 17 und CEN 19 (Tabelle IV.4) macht deutlich, dass die erfolgreiche Produktion von Taxadien von dem richtigen Verhältnis der Enzyme TDS, GGPPS und tHMGR1 zueinander abhängt. In CEN 7 liegen alle Gene einzeln. Sowohl *TDSco*, als auch *GGPPS-Sa* wurden über ein 2µm-basiertes und daher *high-copy*-Plasmid exprimiert, während *tHMGR1* in den einfach pro Zelle vorliegenden Vektor pRS315 ligiert wurde. Dieses Verhältnis der jeweiligen Kopienzahl der Gene führte nicht nur zu der höchsten Taxadienproduktion aller generierten Stämme, sondern auch zur optimalen Relation von Produkt zu Intermediat (Taxadien zu GGOH), bzw. Intermediat zu Nebenprodukt (GGOH zu Squalen, Tabelle IV.4).

**Tabelle IV.4: Vergleich der Produktrelationen in Abhängigkeit der für die Expression verwendeten Vektoren**

Verhältnis	Einzelvektoren (CEN 7)	pVV200 (CEN 17)	YAC (CEN 19)
Taxadien zu GGOH	0,26	0,18	0,14
GGOH zu Squalen	8,48	0,49	0,07

Die Expression aller Gene des Clusters auf dem *high-copy*-Plasmid pVV200 (CEN 17) resultierte zwar unter anderem in einem etwas schlechteren Verhältnis von Taxadien zu GGOH. Offensichtlicher allerdings ist die zehnfache Menge an synthetisiertem Squalen, die ein Intermediat zu Nebenprodukt-Verhältnis bedingt, dass nur einem Siebzehntel dessen von CEN 7 entspricht (Tabelle IV.4). Die Menge an benötigtem tHMGR1 scheint daher geringer zu sein als diejenige von TDS und GGPPS. Sind die Gene in gleicher Kopienzahl vorhanden, sollten die entsprechenden Enzyme in ungefähr gleicher Konzentration vorliegen. Damit werden die durch die Aktivität von tHMGR1 im Überschuss produzierten Vorstufen IPP und DMAPP überwiegend für die Synthese des Nebenprodukts Squalen verwendet und das Verhältnis Produkt zu Nebenprodukt verschlechtert sich. Dass sich dieser Effekt bei hoher Kopienzahl der Gene teilweise relativiert, wird durch den Vergleich von CEN 17 mit CEN 19 deutlich. Bei Expression der fusionierten Gene über das in

einmaliger Kopienzahl pro Zelle vorkommende pYAC (CEN 19) verringert sich das Verhältnis von Intermediat zu Nebenprodukt weiter auf einen Wert, der sich von demjenigen in CEN 7 um zwei Zehnerpotenzen unterscheidet.

Wie bereits in IV.1 aufgeführt, stellt die Taxadiensynthese durch ihre niedrige spezifische Aktivität das Enzym dar, welches für eine optimierte Taxadienproduktion in größter Menge benötigt wird. Da *TDSco* und *GGPPS* bei allen generierten Stämmen in gleicher Kopienzahl vorlagen, wurde bei Analyse der jeweiligen Produkte stets GGOH detektiert, welches auf produziertes GGPP schließen lässt, das nicht zu Taxadien umgesetzt wurde. Obwohl die Verwendung des pYACs für einen *metabolic engineering*-Ansatz zur rekombinanten Produktion eines Sekundärmetaboliten eine Innovation darstellt, muss sie für eine erfolgreiche *de novo*-Synthese von Taxadien oder höher funktionalisierten Taxanen optimiert werden. Dies könnte durch eine aufeinander abgestimmte Expression der benötigten Gene, bzw. Menge entsprechender Enzyme erfolgen, um die Produktion ungewünschter Nebenprodukte zu minimieren. Verschiedene Ansätze wurden bereits entwickelt, die auch für eine rekombinante Produktion von Taxanen Verwendung finden könnten. So erstellte die Gruppe um Stephanopoulos eine Promotorenbibliothek, um die heterologe Expression von Genen genauer dosieren zu können, als es *knock-out* oder Überexpression vermögen.<sup>[141][286]</sup> Die Bibliothek besteht aus mutagenisierten konstitutiven Promotoren, welche durch die Charakterisierung ihrer jeweiligen Stärke eine präzise kontrollierbare Genexpression über einen weiten Bereich verschiedener Expressionslevel ermöglichen. Während diese Methode auf einer Variation in der Expressionsstärke basiert, setzte die Gruppe um Keasling Unterschiede in der posttranskriptionalen Regulation für die Feinabstimmung ein.<sup>[287]</sup> Sie erstellten eine Bibliothek intergener Sequenzen (*Tunable InterGenic Regions*, TIGRs), die aus einer Kombination verschiedener Kontrollelemente bestehen und bei Ligation zwischen die zu exprimierenden Gene Sekundärstrukturen verschiedener Länge ausbilden. Die dadurch bedingte variierende Intensität in Stabilität und Translation synthetisierter mRNS führt zu unterschiedlichen Mengen gebildeten Enzyms. Während bei den Gruppen um Stephanopoulos und Keasling die Feinabstimmung auf Ebene der Transkription und Translation erfolgte, wählten Naesby und Mitarbeiter eine Methode, die nicht ausschließlich von der mRNS- oder Enzymmenge abhängig ist. Dem zufälligen *in vivo*-Zusammenbau der einzelnen Gene eines ganzen Biosynthesewegs in einem YAC schließt sich eine Selektion des Produzenten mit dem höchsten Titer an gewünschtem Endprodukt an. Der Ansatz ist daher unabhängig von rationalen Designaspekten des Biosynthesewegs auf Transkriptions-, Translations- oder Aktivitätsebene.

Auch für die optimierte Synthese von Taxadien wäre ein zufälliger *in vivo*-Zusammenbau der Gene denkbar. Durch die Selektion des erfolgreichsten Taxadienproduzenten mittels GC/MS wäre die jeweils benötigte Menge der einzelnen Enzyme des Stamms aufeinander abgestimmt. Die entsprechende Kopienzahl der einzelnen Gene könnte anschließend durch Reextraktion des konstruierten YACs und seiner Sequenzierung ermittelt werden. Weiterhin wäre über diese Methode die gleichzeitige Insertion nachfolgender Taxolbiosynthesegene möglich. Aufgrund der steigenden Anzahl funktioneller Gruppen (Abbildung I.17), müsste die Selektion entsprechender Produzenten hierbei allerdings mithilfe der LC/MS-Analytik erfolgen.

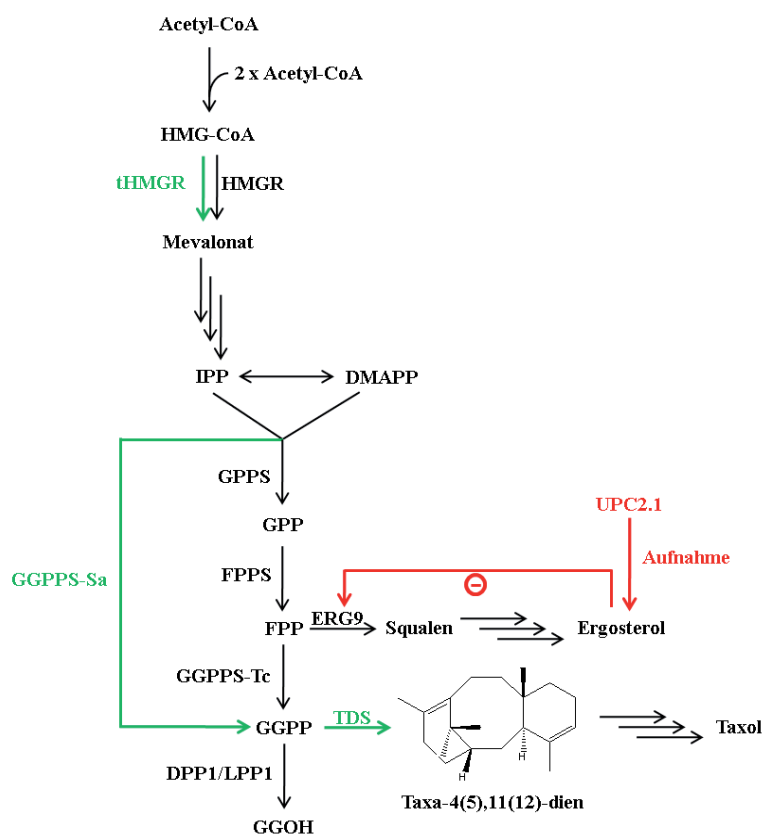
Die momentane Quelle von Taxol und seinen Analoga besteht aus der Semisynthese ausgehend von aus der Pflanze oder Zellkultur isolierten Intermediaten. Um diese durch eine rekombinante Synthese mithilfe von Mikroorganismen zu ersetzen, müssen neben den hier verwendeten Genen zur Biosynthese von Taxadien zahlreiche zusätzliche Synthesegene exprimiert werden. Mittlerweile wurden allerdings verschiedene Methoden entwickelt, die eine Generierung ausbalancierter Cluster ermöglichen (Tabelle IV.2). Die dauerhafte Stabilität des generierten pYACs als künstliches Hefechromosom muss durch eine Linearisierung zur Freisetzung der Telomere noch untersucht werden. Dennoch stellen die generierten Hefestämme CEN 18 und CEN 19 durch die erstmalige Fusionierung aller für die optimierte Biosynthese von Taxadien benötigten Gene eine brauchbare Plattform für die *de novo*-Synthese höher funktionalisierter Taxane dar.

## V Zusammenfassung

Taxol ist ein terpenoider Naturstoff, der im Rahmen einer weltweiten Suche nach biologisch aktiven Substanzen in der Eibenrinde entdeckt wurde. Aufgrund seiner antikanzerogenen Wirkung und seines speziellen Wirkungsmechanismus wird er heute weltweit als Medikament gegen zahlreiche Krebsarten eingesetzt. Die Extraktion der spärlich synthetisierten Mengen aus den natürlichen Quellen *Taxus* spp. und verschiedenen Endophyten kann den Bedarf nicht decken – ebenso wenig die entwickelten Verfahren zur Totalsynthese, die wegen der komplexen Molekülstruktur noch zu viele Reaktionsschritte mit geringer Gesamtausbeute benötigen. Die Gewinnung von Taxol basiert daher zum Einen auf der Isolierung der in größeren Mengen biosynthetisierten Vorstufen Baccatin III oder 10-Deacetylbaccatin III, die anschließend chemisch weiter zu den entsprechenden Generika Taxotere oder Paclitaxel umgesetzt werden. Zum anderen wird Taxol aus Zellkulturen extrahiert. Es existieren bisher keine Methoden zur stabilen genetischen Veränderung von *Taxus* spp., sodass eine Optimierung der Taxolbiosynthese ausschließlich über die Verbesserung von Fermentationsbedingungen der Pflanzenzellkulturen erfolgen kann und dementsprechend begrenzt ist. Der Einsatz rekombinanter Mikroorganismen zur Produktion der Verbindungen verkörpert daher eine prinzipiell zeit- und kostengünstigere Methode der Synthese. Zusätzlich ist Taxol in seiner natürlichen Quelle nur eins von hunderten gebildeter Taxanderivate, welches nur einen geringen Anteil der gesamten Taxanfraktion ausmacht. Die Aufreinigung der gewünschten Verbindung aus einem Mikroorganismus, der ausschließlich die rekombinanten Gene für die Biosynthese von Taxol aufweist, würde daher kaum Probleme hinsichtlich der Abtrennung ähnlicher Substanzen involvieren. Im Rahmen dieser Arbeit wurde *S. cerevisiae* wegen seines Status als eukaryotischer Modellorganismus und seines erfolgreichen Einsatzes bei bisherigen Untersuchungen und Charakterisierungen von Taxolbiosynthesegenen gewählt.

Ziel war es, einen Hefestamm zu generieren, der basierend auf einer optimierten *de novo*-Synthese von Taxadien als dem ersten Intermediat der Taxolbiosynthese fähig ist, höher funktionalisierte Taxane in quantifizierbaren Mengen zu produzieren. Die Optimierung sollte anhand einer heterologen Expression von *GGPPS* und *TDS* realisiert werden, sowie der zusätzlichen Modifizierung des hefeeigenen Primärstoffwechsels. Zu diesem Zweck wurden verschiedene Gene auf die Auswirkung ihrer Expression auf eine Akkumulation der diterpenoiden Vorstufe GGPP, bzw. des Diterpens Taxadien in *S. cerevisiae* hin untersucht. Einen der analysierten Modifikationsansätze verkörpert die Deregulierung des Mevalonatwegs für die effiziente Bereitstellung der Endprodukte IPP und DMAPP. Da die Regulation in der Umsetzung von HMG-CoA zu Mevalonat durch die HMG-CoA-Reduktase erfolgt, wurde eine verkürzte Version des

hefeeigenen HMG-CoA-Reduktasesens 1 verwendet, dessen Produkt ausschließlich aus einer katalytischen Domäne besteht und nicht mehr reguliert ist. Weiterhin wurde der Kohlenstofffluss umgeleitet. FPP stellt einen metabolischen Verzweigungspunkt dar, an dem sich die Biosynthese von Sesqui- und Triterpenen von derjenigen GGPP-abgeleiteter Verbindungen trennt (Abbildung V.1). Um die Verwendung isoprenoider Vorstufen für die Synthese unerwünschter FPP-abgeleiteter Substanzen zu umgehen, wurde *GGPPS-Sa* exprimiert. Dieses archaebakterielle Gen kodiert für ein bifunktionales Enzym, welches die multiple Kondensation von IPP und DMAPP bis hin zu GGPP katalysiert. Die Expression eines punktmutierten Transkriptionsfaktors (*upc2.1*) sollte zusätzlich die Sterolaufnahme aus dem Medium unter aeroben Bedingungen ermöglichen, sodass die gebildeten Isoprenoidvorstufen nicht in die hefeeigene Sterolbiosynthese geleitet werden. Erst die Kombination aller Ansätze mit der gleichzeitigen Verwendung eines für die Expression in *S. cerevisiae* codon-optimierten Taxadiensynthasegens zur Umgehung der beobachteten Translationsprobleme führte zu einer Taxadienproduktion, die mit über 8 mg/L die bisher bekannten Syntheseleistungen um eine Zehnerpotenz übersteigt. Abbildung V.1 gibt die Modifizierungen des in dieser Arbeit generierten Taxadien-produzierenden Hefestamms CEN 7 schematisch wieder.



**Abbildung V.1: Schematische Darstellung der verwendeten Modifizierungsansätze zur Generierung eines Taxadien-produzierenden Hefestamms.** HMG-CoA: 3-Hydroxy-3-methylglutaryl-CoA; HMGR: HMG-CoA-Reduktase; tHMGR: um die regulatorische Domäne verkürzte Version der hefeeigenen HMG-CoA-Reduktase zur Deregulation des Mevalonatswegs und damit ausreichenden Bereitstellung der isoprenoiden Vorstufen IPP und DMAPP; IPP: Isopentenylidiphosphat; DMAPP: Dimethylallyldiphosphat; GPP: Geranylidiphosphat; GPPS: GPP-Synthase; FPP: Farnesylidiphosphat; FPPS: FPP-Synthase; ERG9: Squalensynthase; GGPP: Geranylgeranylidiphosphat; GGPPS-Tc: GGPP-Synthase aus *T. chinensis*; GGPPS-Sa: GGPP-Synthase aus *S. acidocaldarius* katalysiert als bifunktionales Enzym die Kondensation von IPP und DMAPP über GPP und FPP bis hin zum GGPP; TDS: Taxadiensynthase; DPP1/LPP1: Phosphatasen, die GGPP (parallel zur Zyklisierung zu Taxadien durch die TDS) zu GGOH dephosphorylieren; GGOH: Geranylgeraniol; UPC2.1: punktmutierte Version des Transkriptionsfaktors UPC2, ermöglicht bei Expression die Aufnahme von Sterol auch unter aeroben Bedingungen; Grün/Rot: Heterologes Enzyme

phorylieren; GGOH: Geranylgeraniol; UPC2.1: punktmutierte Version des Transkriptionsfaktors UPC2, ermöglicht bei Expression die Aufnahme von Sterol auch unter aeroben Bedingungen; Grün/Rot: Heterologes Enzyme



Um trotz der limitierten Anzahl an Selektionsmarkern weitere Gene der Taxolbiosynthese für die rekombinante Produktion von funktionalisierten Taxanen exprimieren zu können, sollten die für eine optimierte Taxadiensynthese benötigten Gene fusioniert und auf einem Plasmid koexprimiert werden. Dazu wurden *TDSco*, *GGPPS-Sa* und *tHMGR* jeweils mit einem individuellen Promotor und Terminator fusioniert und durch Transformation des Hefestamms CEN.PK2-1C mit dem entstandenen Gencluster der Fusionsstamm CEN 8 generiert. Die Analyse akkumulierter Produkte mittels GC/MS, sowie Northern- und Westernblot ließen auf ein nicht funktionelles Fusionskonstrukt schließen. Um die Orientierung der Gene zueinander oder die Wahl der verwendeten Promotor-, bzw. Terminatorregion als Ursache auszuschließen, wurden neun Fusionskonstrukte generiert, die in diesen Parametern variierten. Trotz korrekter Basenabfolge führte nach Transformation keines der Cluster zu einem Hefestamm, in dem alle Gene funktionell exprimiert wurden. Eine genauere Untersuchung der Fusionskonstrukte demonstrierte schließlich ihre Veränderung innerhalb der Hefezelle. Durch Reextraktion der einzelnen Cluster aus dem jeweiligen Stamm zu verschiedenen Zeitpunkten wurde weiterhin deutlich, dass diese Veränderung unabhängig vom Konstrukt und der Inkubationszeit quantitativ, aber nicht einheitlich erfolgte. Der für die Expression der Cluster verwendete Vektor veränderte sich nachweislich sowohl als Leervektor, als auch bei Expression eines variablen Inserts in der Hefezelle. Der Grund für die beobachtete Veränderung wurde daher im Vektorgrundgerüst vermutet und das Cluster zur Umgehung des Problems mithilfe eines anderen Plasmids exprimiert. Die Transformation führte zu einem Fusionsstamm (CEN 18), der nicht nur alle drei Gene transkribierte und die entsprechenden rekombinanten Proteine synthetisierte, sondern ebenfalls Taxadien im mg-Maßstab produzierte. Die anschließende Expression des Clusters mithilfe eines zirkulären YACs (CEN 19) konnte ebenfalls erfolgreich durchgeführt werden.

Zusammengefasst wurde mit CEN 7 ein Hefestamm generiert, der in der Lage ist, die bisher größte bekannte Menge an Taxadien zu produzieren. Das ursprüngliche Ziel der Synthese eines funktionalisierten Taxans wurde hingegen nicht erreicht. Dies lag in der problematischen Konstruktion des Clusters begründet, welches für eine Selektion zusätzlicher Taxolbiosynthesegene nötig ist. Die aufgetretenen Schwierigkeiten konnten allerdings auf die Instabilität des verwendeten Vektorgrundgerüsts pAG423 zurückgeführt werden. Durch den Austausch des pAG423- gegen den pVV200-Vektor gelang es nicht nur, das Problem zu lösen, sondern weiterhin wurde erstmalig ein rekombinanter Hefestamm generiert, der in der Lage ist, mithilfe eines Genclusters Taxadien herzustellen. Somit wurde eine Plattform entwickelt, die mittels Expression weiterer Taxolbiosynthesegene die Biosynthese funktionalisierter Taxane ermöglichen sollte. Die



Verwendung eines pYACs für die Expression des Clusters muss für eine tragfähige *de novo*-Synthese von Taxadien oder funktionalisierten Taxanen zwar weiter optimiert werden, verkörpert aber einen erfolgreichen, innovativen Ansatz des *metabolic engineerings* zur rekombinanten Produktion von Sekundärmetaboliten.

## VI Referenzen

- [1] P. K. Ajikumar, K. Tyo, S. Carlsen, O. Mucha, T. H. Phon, G. Stephanopoulos, *Molecular Pharmaceutics* **2008**, 5, 167.
- [2] E. Breitmaier, *Terpene - Aromen, Düfte, Pharmaka, Pheromone*, Wiley VCH Verlag GmbH, **2005**.
- [3] L. Ruzicka, *Proceedings of the Chemical Society of London* **1959**, 341.
- [4] a) J. Bohlmann, J. Crock, R. Jetter, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **1998**, 95, 6756; b) S. C. Trapp, R. B. Croteau, *Genetics* **2001**, 158, 811.
- [5] A. Aharoni, A. P. Giri, F. W. A. Verstappen, C. M. Berteau, R. Sevenier, Z. Sun, M. A. Jongsma, W. Schwab, H. J. Bouwmeester, *Plant Cell* **2004**, 16, 3110.
- [6] D. J. Mcgarvey, R. Croteau, *Plant Cell* **1995**, 7, 1015.
- [7] R. B. Croteau, E. M. Davis, K. L. Ringer, M. R. Wildung, *Naturwissenschaften* **2005**, 92, 562.
- [8] A. K. Singh, V. K. Raina, A. A. Naqvi, N. K. Patra, B. Kumar, P. Ram, S. P. S. Khanuja, *Flavour and Fragrance Journal* **2005**, 20, 302.
- [9] R. Croteau, *Planta Medica* **1991**, 57, S10.
- [10] D. L. Klayman, *Science* **1985**, 228, 1049.
- [11] S. Yeung, W. Pongtavornpinyo, I. M. Hastings, A. J. Mills, N. J. White, *American Journal of Tropical Medicine and Hygiene* **2004**, 71, 179.
- [12] T. E. Wallaart, N. Pras, A. C. Beekman, W. J. Quax, *Planta Medica* **2000**, 66, 57.
- [13] S. Kumar, S. Srivastava, *Current Science* **2005**, 89, 1097.
- [14] P. S. Covello, *Phytochemistry* **2008**, 69, 2881.
- [15] K. Reiling, N. McPhee, K. Fisher, D. Ockey, **2006**.
- [16] H. J. Bouwmeester, T. E. Wallaart, M. H. A. Janssen, B. van Loo, B. J. M. Jansen, M. A. Posthumus, C. O. Schmidt, J. W. De Kraker, W. A. König, M. C. R. Franssen, *Phytochemistry* **1999**, 52, 843.
- [17] T. E. Wallaart, H. J. Bouwmeester, J. Hille, L. Poppinga, N. C. A. Maijers, *Planta* **2001**, 212, 460.
- [18] M. Suffness, E. Wall, **1995**.
- [19] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggon, A. T. Mcphail, *Journal of the American Chemical Society* **1971**, 93, 2325.
- [20] M. E. Wall, *Medicinal Research Reviews* **1998**, 18, 299.
- [21] M. Suffness, *Taxane Anticancer Agents* **1995**, 583, 1.
- [22] D. G. I. Kingston, *Abstracts of Papers of the American Chemical Society* **1992**, 203, 111.
- [23] D. G. I. Kingston, *Phytochemistry* **2007**, 68, 1844.
- [24] P. B. Schiff, J. Fant, S. B. Horwitz, *Nature* **1979**, 277, 665.
- [25] V. Bartsch, *Angriff auf die mitotische Spindel: Wirkungsmechanismus der Taxane*, Vol. 2, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Weinheim, **2005**.
- [26] E. Hamel, D. L. Sackett, D. Vourloumis, K. C. Nicolaou, *Biochemistry* **1999**, 38, 5490.
- [27] O. Exposito, M. Bonfill, E. Moyano, M. Onrubia, M. H. Mirjalili, R. M. Cusido, J. Palazon, *Anti-Cancer Agents in Medicinal Chemistry* **2009**, 9, 109.

- [28] A. Farb, P. F. Heller, S. Shroff, L. Cheng, F. D. Kolodgie, A. J. Carter, D. S. Scott, J. Froehlich, R. Virmani, *Circulation* **2001**, *104*, 473.
- [29] J. M. Lasala, G. W. Stone, K. D. Dawkins, P. W. Serruys, A. Colombo, E. Grube, J. Koglin, S. Ellis, *Journal of Interventional Cardiology* **2006**, *19*, 422.
- [30] R. B. Weiss, R. C. Donehower, P. H. Wiernik, T. Ohnuma, R. J. Gralla, D. L. Trump, J. R. Baker, D. A. Vanecho, D. D. Vonhoff, B. Leylandjones, *Journal of Clinical Oncology* **1990**, *8*, 1263.
- [31] D. G. I. Kingston, *Trends in Biotechnology* **1994**, *12*, 222.
- [32] D. Guenard, F. Guerittevoegelein, P. Potier, *Accounts of Chemical Research* **1993**, *26*, 160.
- [33] G. Appendino, P. Gariboldi, B. Gabetta, R. Pace, E. Bombardelli, D. Viterbo, *Journal of the Chemical Society, Perkin Transactions 1* **1992**, 2925.
- [34] M. Distefano, G. Scambia, C. Ferlini, C. Gaggini, R. D. Vincenzo, A. Riva, E. Bombardelli, I. Ojima, A. Fattorossi, P. B. Panici, S. Mancuso, *International Journal of Cancer* **1997**, *72*, 844.
- [35] M. I. Nicoletti, T. Colombo, C. Rossi, C. Monardo, S. Stura, M. Zucchetti, A. Riva, P. Morazzoni, M. B. Donati, E. Bombardelli, M. D'Incalci, R. Giavazzi, *Cancer Research* **2000**, *60*, 842.
- [36] L. Barboni, G. Giarlo, R. Ballini, G. Fontana, *Bioorganic & Medicinal Chemistry Letters* **2006**, *16*, 5389.
- [37] N. C. Wheeler, K. Jech, S. Masters, S. W. Brobst, A. B. Alvarado, A. J. Hoover, K. M. Snader, *Journal of Natural Products* **1992**, *55*, 432.
- [38] M. Nadeem, H. C. Rikhari, A. Kumar, L. M. S. Palni, S. K. Nandi, *Phytochemistry* **2002**, *60*, 627.
- [39] G. Nimachow, J. S. Rawat, O. Dai, *Current Science* **2010**, *98*, 1434.
- [40] A. Stierle, G. Strobel, D. Stierle, *Science* **1993**, *260*, 214.
- [41] G. Strobel, X. S. Yang, J. Sears, R. Kramer, R. S. Sidhu, W. M. Hess, *Microbiology-Uk* **1996**, *142*, 435.
- [42] G. Kathiravan, V. Sri Raman, *Fitoterapia* **2010**, *81*, 557.
- [43] R. A. Holton, C. Somoza, H. B. Kim, F. Liang, R. J. Biediger, P. D. Boatman, M. Shindo, C. C. Smith, S. C. Kim, H. Nadizadeh, Y. Suzuki, C. L. Tao, P. Vu, S. H. Tang, P. S. Zhang, K. K. Murthi, L. N. Gentile, J. H. Liu, *Journal of the American Chemical Society* **1994**, *116*, 1597.
- [44] R. A. Holton, H. B. Kim, C. Somoza, F. Liang, R. J. Biediger, P. D. Boatman, M. Shindo, C. C. Smith, S. C. Kim, H. Nadizadeh, Y. Suzuki, C. L. Tao, P. Vu, S. H. Tang, P. S. Zhang, K. K. Murthi, L. N. Gentile, J. H. Liu, *Journal of the American Chemical Society* **1994**, *116*, 1599.
- [45] K. C. Nicolaou, Z. Yang, J. J. Liu, H. Ueno, P. G. Nantermet, R. K. Guy, C. F. Claiborne, J. Renaud, E. A. Couladouros, K. Paulvannan, E. J. Sorensen, *Nature* **1994**, *367*, 630.
- [46] T. Doi, S. Fuse, S. Miyamoto, K. Nakai, D. Sasuga, T. Takahashi, *Chemistry - An Asian Journal* **2006**, *1*, 370.
- [47] a) T. Mukaiyama, I. Shiina, H. Iwadare, M. Saitoh, T. Nishimura, N. Ohkawa, H. Sakoh, K. Nishimura, Y.-I. Tani, M. Hasegawa, K. Yamada, K. Saitoh, *Chemistry - An European Journal* **1999**, *5*, 121; b) I. Shiina, K. Saitoh, I. Frechard-Ortuno, T. Mukaiyama, *Chemical Letters* **1998**, 3.

- [48] S. J. Danishefsky, J. J. Masters, W. B. Young, J. T. Link, L. B. Snyder, T. V. Magee, D. K. Jung, R. C. A. Isaacs, W. G. Bornmann, C. A. Alaimo, C. A. Coburn, M. J. Di Grandi, *Journal of the American Chemical Society* **1996**, *118*, 2843.
- [49] P. A. Wender, N. F. Badham, S. P. Conway, P. E. Floreancig, T. E. Glass, J. B. Houze, N. E. Krauss, D. Lee, D. G. Marquess, P. L. McGrane, W. Meng, M. G. Natchus, A. J. Shuker, J. C. Sutton, R. E. Taylor, *Journal of the American Chemical Society* **1997**, *119*, 2757.
- [50] J. N. Denis, A. E. Greene, D. Guenard, F. Guerittevoegelein, L. Mangatal, P. Potier, *Journal of the American Chemical Society* **1988**, *110*, 5917.
- [51] K. M. Witherup, S. A. Look, M. W. Stasko, T. J. Ghiorzi, G. M. Muschik, G. M. Cragg, *Journal of Natural Products* **1990**, *53*, 1249.
- [52] S. Ritter, *Chemical and Engineering News* **2004**, *82*, 25.
- [53] D. M. Gibson, R. E. B. Ketchum, N. C. Vance, A. A. Christen, *Plant Cell Reports* **1993**, *12*, 479.
- [54] A. G. Fett-Neto, F. DiCosmo, W. F. Reynolds, K. Sakata, *Nature Biotechnology* **1992**, *10*, 1572.
- [55] E. R. M. Wickremesinhe, R. N. Arteca, *Plant Cell Tissue and Organ Culture* **1993**, *35*, 181.
- [56] V. Srinivasan, L. Pestchanker, S. Moser, T. J. Hirasuna, R. A. Taticek, M. L. Shuler, *Biotechnology and Bioengineering* **1995**, *47*, 666.
- [57] Y. Yukimune, H. Tabata, Y. Higashi, Y. Hara, *Nature Biotechnology* **1996**, *14*, 1129.
- [58] R. E. B. Ketchum, J. V. Luong, D. M. Gibson, *Journal of Liquid Chromatography & Related Technologies* **1999**, *22*, 1715.
- [59] H. Tabata, in *Biomanufacturing, Vol. 87*, Springer Berlin / Heidelberg, **2004**, pp. 1.
- [60] R. A. Holton, in *Taxol: Science and application* (Ed.: M. Suffness), CRC Press, **1995**, pp. 97.
- [61] F. Lynen, J. Knappe, H. Eggerer, U. Henning, B. W. Agranoff, *Federation Proceedings* **1959**, *18*, 278.
- [62] R. D. Tanaka, B. L. Schafer, L. Y. Lee, J. S. Freudenberger, S. T. Mosley, *Journal of Biological Chemistry* **1990**, *265*, 2391.
- [63] D. Dimsterdenk, M. K. Thorsness, J. Rine, *Molecular Biology of the Cell* **1994**, *5*, 655.
- [64] M. E. Basson, M. Thorsness, J. Rine, *Proceedings of the National Academy of Sciences of the United States of America* **1986**, *83*, 5563.
- [65] R. Y. Hampton, J. Rine, *Journal of Cellular Biology* **1994**, *125*, 299.
- [66] J. L. Goldstein, M. S. Brown, *Nature (London)* **1990**, *343*, 425.
- [67] R. Hampton, D. DimsterDenk, J. Rine, *Trends in Biochemical Sciences* **1996**, *21*, 140.
- [68] M. Basson, J. Rine, *Journal of Lipid Research* **1983**, *24*, 1405.
- [69] K. A. Donald, R. Y. Hampton, I. B. Fritz, *Applied and Environmental Microbiology* **1997**, *63*, 3341.
- [70] M. Thorsness, W. Schafer, L. D'Ari, J. Rine, *Molecular Cell Biology* **1989**, *9*, 5702.
- [71] C. L. Ward, S. Omura, R. R. Kopito, *Cell* **1995**, *83*, 121.
- [72] R. Y. Hampton, H. Bhakta, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 12944.
- [73] R. Y. Hampton, R. M. Garza, *Chemical Reviews* **2009**, *109*, 1561.
- [74] S. Spurgeon, J. Porter, in *Biosynthesis of Isoprenoid Compounds* (Eds.: S. Spurgeon, J. Porter), John Wiley & Sons, New York, **1981**, pp. 1.

- [75] D. Zhou, R. H. White, *Biochemical Journal* **1991**, 273, 627.
- [76] T. S. Raman, B. V. S. Sharma, Jayarama.J, Ramasarm.T, *Archives of Biochemistry and Biophysics* **1965**, 110, 75.
- [77] M. Rohmer, M. Knani, P. Simonin, B. Sutter, H. Sahm, *Biochemical Journal* **1993**, 295, 517.
- [78] M. Rohmer, M. Seemann, S. Horbach, S. BringerMeyer, H. Sahm, *Journal of the American Chemical Society* **1996**, 118, 2564.
- [79] W. Eisenreich, M. Schwarz, A. Cartayrade, D. Arigoni, M. H. Zenk, A. Bacher, *Chemistry & Biology* **1998**, 5, R221.
- [80] D. Arigoni, S. Sagner, C. Latzel, W. Eisenreich, A. Bacher, M. H. Zenk, *Proceedings of the National Academy of Sciences of the United States of America* **1997**, 94, 10600.
- [81] W. Eisenreich, B. Menhard, P. J. Hylands, M. H. Zenk, A. Bacher, *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, 6431.
- [82] A. E. Koepp, M. Hezari, J. Zajicek, B. S. Vogel, R. E. Lafever, N. G. Lewis, R. Croteau, *Journal of Biological Chemistry* **1995**, 270, 8686.
- [83] M. Hezari, A. Koepp, R. Croteau, *Abstracts of Papers of the American Chemical Society* **1995**, 210, 172.
- [84] X. Y. Lin, M. Hezari, A. E. Koepp, H. G. Floss, R. Croteau, *Biochemistry* **1996**, 35, 2968.
- [85] F. Guerittevoegelein, D. Guenard, P. Potier, *Journal of Natural Products* **1987**, 50, 9.
- [86] D. C. Williams, B. J. Carroll, Q. Jin, C. D. Rithner, S. R. Lenger, H. G. Floss, R. M. Coates, R. M. Williams, R. Croteau, *Chemical Biology* **2000**, 7, 969.
- [87] Q. W. Jin, D. C. Williams, M. Hezari, R. Croteau, R. M. Coates, *Journal of Organic Chemistry* **2005**, 70, 4667.
- [88] Y. H. Jin, D. C. Williams, R. Croteau, R. M. Coates, *Journal of the American Chemical Society* **2005**, 127, 7834.
- [89] M. R. Wildung, R. Croteau, *Journal of Biological Chemistry* **1996**, 271, 9201.
- [90] G. Vonheijne, J. Steppuhn, R. G. Herrmann, *European Journal of Biochemistry* **1989**, 180, 535.
- [91] D. C. Williams, M. R. Wildung, A. Q. W. Jin, D. Dalal, J. S. Oliver, R. M. Coates, R. Croteau, *Archives of Biochemistry and Biophysics* **2000**, 379, 137.
- [92] G. P. Bolwell, K. Bozak, A. Zimmerlin, *Phytochemistry* **1994**, 37, 1491.
- [93] U. Heinig, S. Jennewein, *African Journal of Biotechnology* **2009**, 8, 1370.
- [94] R. Croteau, R. Ketchum, R. Long, R. Kaspera, M. Wildung, *Phytochemistry Reviews* **2006**, 5, 75.
- [95] C. Chapple, *Annual Review of Plant Physiology and Plant Molecular Biology* **1998**, 49, 311.
- [96] M. A. Schuler, D. Werck-Reichhart, *Annual Review of Plant Biology* **2003**, 54, 629.
- [97] K. S. Ahn, E. Szczesnaskorupa, B. Kemper, *Journal of Biological Chemistry* **1993**, 268, 18726.
- [98] J. Hefner, S. M. Rubenstein, R. E. B. Ketchum, D. M. Gibson, R. M. Williams, R. Croteau, *Chemistry & Biology* **1996**, 3, 479.
- [99] S. Jennewein, R. M. Long, R. M. Williams, R. Croteau, *Chemistry & Biology* **2004**, 11, 379.
- [100] K. Walker, A. Schoendorf, R. Croteau, *Archives of Biochemistry and Biophysics* **2000**, 374, 371.



- [101] A. Schoendorf, C. D. Rithner, R. M. Williams, R. B. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 1501.
- [102] J. Hefner, R. E. B. Ketchum, R. Croteau, *Archives of Biochemistry and Biophysics* **1998**, 360, 62.
- [103] A. L. Wheeler, R. M. Long, R. E. B. Ketchum, C. D. Rithner, R. M. Williams, R. Croteau, *Archives of Biochemistry and Biophysics* **2001**, 390, 265.
- [104] S. Jennewein, C. D. Rithner, R. M. Williams, R. B. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2001**, 98, 13595.
- [105] S. Jennewein, C. D. Rithner, R. M. Williams, R. Croteau, *Archives of Biochemistry and Biophysics* **2003**, 413, 262.
- [106] E. Baloglu, D. G. I. Kingston, *Journal of Natural Products* **1999**, 62, 1448.
- [107] D. Hampel, C. J. D. Mau, R. B. Croteau, *Archives of Biochemistry and Biophysics* **2009**, 487, 91.
- [108] R. E. B. Ketchum, T. Horiguchi, D. Y. Qiu, R. M. Williams, R. B. Croteau, *Phytochemistry* **2007**, 68, 335.
- [109] D. Pompon, B. Louerat, A. Bronine, P. Urban, *Cytochrome P450, Pt B* **1996**, 272, 51.
- [110] M. Chau, S. Jennewein, K. Walker, R. Croteau, *Chemistry & Biology* **2004**, 11, 663.
- [111] M. Chau, R. Croteau, *Archives of Biochemistry and Biophysics* **2004**, 427, 48.
- [112] K. Walker, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 583.
- [113] K. Walker, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 13591.
- [114] J. Parness, D. G. I. Kingston, R. G. Powell, C. Harracksingh, S. B. Horwitz, *Biochemical and Biophysical Research Communications* **1982**, 105, 1082.
- [115] K. Walker, S. Fujisaki, R. Long, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 12715.
- [116] K. Walker, R. Long, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 9166.
- [117] R. M. Long, R. Croteau, *Biochemical and Biophysical Research Communications* **2005**, 338, 410.
- [118] R. M. Long, C. Lagisetti, R. M. Coates, R. B. Croteau, *Archives of Biochemistry and Biophysics* **2008**, 477, 384.
- [119] S. Jennewein, M. R. Wildung, M. Chau, K. Walker, R. Croteau, *Proceedings of the National Academy of Sciences of the United States of America* **2004**, 101, 9149.
- [120] J. E. Bailey, *Science* **1991**, 252, 1668.
- [121] C. D. Smolke, *The metabolic pathway engineering handbook* **2010**, CRC Press New York
- [122] R. Barkovich, J. C. Liao, *Metabolic Engineering* **2001**, 3, 27.
- [123] N. Misawa, M. Nakagawa, K. Kobayashi, S. Yamano, Y. Izawa, K. Nakamura, K. Harashima, *Journal of Bacteriology* **1990**, 172, 6704.
- [124] S. Yamano, T. Ishii, M. Nakagawa, H. Ikenaga, N. Misawa, *Bioscience Biotechnology and Biochemistry* **1994**, 58, 1112.
- [125] Y. Miura, K. Kondo, T. Saito, H. Shimada, P. D. Fraser, N. Misawa, *Applied and Environmental Microbiology* **1998**, 64, 1226.
- [126] B. E. Jackson, E. A. Hart-Wells, S. P. T. Matsuda, *Organic Letters* **2003**, 5, 1629.

- [127] C. W. Wang, M. K. Oh, J. C. Liao, *Biotechnology and Bioengineering* **1999**, 62, 235.
- [128] C. w. Wang, M. K. Oh, J. C. Liao, *Biotechnology Progress* **2000**, 16, 922.
- [129] S. Kajiwara, P. D. Fraser, K. Kondo, N. Misawa, *Biochemical Journal* **1997**, 324, 421.
- [130] S. W. Kim, J. D. Keasling, *Biotechnology and Bioengineering* **2001**, 72, 408.
- [131] P. D. Matthews, E. T. Wurtzel, *Applied Microbiology and Biotechnology* **2000**, 53, 396.
- [132] M. Albrecht, N. Misawa, G. Sandmann, *Biotechnology Letters* **1999**, 21, 791.
- [133] L. Z. Yuan, P. E. Rouviere, R. A. LaRossa, W. Suh, *Metabolic Engineering* **2006**, 8, 79.
- [134] V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nature Biotechnology* **2003**, 21, 796.
- [135] V. J. J. Martin, Y. Yoshikuni, J. D. Keasling, *Biotechnology and Bioengineering* **2001**, 75, 497.
- [136] W. R. Farmer, J. C. Liao, *Biotechnology Progress* **2001**, 17, 57.
- [137] R. V. Vadali, Y. C. Fu, G. N. Bennett, K. Y. San, *Biotechnology Progress* **2005**, 21, 1558.
- [138] M. Takagi, T. Kuzuyama, S. Takahashi, H. Seto, *Journal of Bacteriology* **2000**, 182, 4153.
- [139] Y. T. Yang, G. N. Bennett, K. Y. San, *Biotechnology and Bioengineering* **1999**, 65, 291.
- [140] H. Alper, Y. S. Jin, J. F. Moxley, G. Stephanopoulos, *Metabolic Engineering* **2005**, 7, 155.
- [141] H. Alper, K. Miyaoku, G. Stephanopoulos, *Nature Biotechnology* **2005**, 23, 612.
- [142] Q. L. Huang, C. A. Roessner, R. Croteau, A. I. Scott, *Bioorganic & Medicinal Chemistry* **2001**, 9, 2237.
- [143] T. Kuzuyama, M. Takagi, S. Takahashi, H. Seto, *Journal of Bacteriology* **2000**, 182, 891.
- [144] M. Hotze, G. Schroder, J. Schroder, *FEBS Letters* **1995**, 374, 345.
- [145] G. Schröder, E. Unterbusch, M. Kaltenbach, J. Schmidt, D. Strack, V. De Luca, J. Schröder, *FEBS Letters* **1999**, 458, 97.
- [146] O. A. Carter, R. J. Peters, R. Croteau, *Phytochemistry* **2003**, 64, 425.
- [147] O. Besumbes, S. Sauret-Gueto, M. A. Phillips, S. Imperial, M. Rodriguez-Concepcion, A. Boronat, *Biotechnology and Bioengineering* **2004**, 88, 168.
- [148] A. Anterola, E. Shanle, P. F. Perroud, R. Quatrano, *Transgenic Research* **2009**, 18, 655.
- [149] J. M. DeJong, Y. L. Liu, A. P. Bollon, R. M. Long, S. Jennewein, D. Williams, R. B. Croteau, *Biotechnology and Bioengineering* **2006**, 93, 212.
- [150] S. Jennewein, H. Park, J. M. DeJong, R. M. Long, A. P. Bollon, R. B. Croteau, *Biotechnology and Bioengineering* **2005**, 89, 588.
- [151] S. P. S. Khanuja, S. Paul, A. K. Shasany, A. K. Gupta, M. P. Darokar, M. M. Gupta, R. K. Verma, G. Ram, A. Kumar, R. K. Lal, R. P. Bansal, A. K. Singh, R. S. Bhakuni, S. Tandon, (Council of Scientific & Industrial Research, India). Application: US patent **2005**
- [152] C. Z. Liu, Y. Zhao, Y. C. Wang, *Applied Microbiology and Biotechnology* **2006**, 72, 11.
- [153] J. P. van Dijken, J. Bauer, L. Brambilla, P. Duboc, J. M. Francois, C. Gancedo, M. L. Giuseppin, J. J. Heijnen, M. Hoare, H. C. Lange, E. A. Madden, P. Niederberger, J. Nielsen, J. L. Parrou, T. Petit, D. Porro, M. Reuss, N. van Riel, M. Rizzi, H. Y. Steensma, C. T. Verrips, J. Vindelov, J. T. Pronk, *Enzyme and Microbial Technology* **2000**, 26, 706.
- [154] V. Van Mullem, M. Wery, X. De Bolle, J. Vandenhaute, *Yeast* **2003**, 20, 739.
- [155] R. S. Sikorski, P. Hieter, *Genetics* **1989**, 122, 19.
- [156] S. Alberti, A. D. Gitler, S. Lindquist, *Yeast* **2007**, 24, 913.
- [157] E. Gari, L. Piedrafita, M. Aldea, E. Herrero, *Yeast* **1997**, 13, 837.
- [158] D. T. Burke, G. F. Carle, M. V. Olson, *Science* **1987**, 236, 806.



- [159] M. A. U. Mannan, S. Sharma, K. Ganesan, *Analytical Biochemistry* **2009**, 389, 77.
- [160] R. M. Horton, H. D. Hunt, S. N. Ho, J. K. Pullen, L. R. Pease, *Gene* **1989**, 77, 61.
- [161] F. Sanger, J. E. Donelson, A. R. Coulson, H. Kossel, D. Fischer, *Journal of Molecular Biology* **1974**, 90, 315.
- [162] C. J. Watson, J. E. Froehlich, C. A. Josefsson, C. Chapple, F. Durst, I. Benveniste, R. C. Coolbaugh, *Plant and Cell Physiology* **2001**, 42, 873.
- [163] D. R. Nelson, H. W. Strobel, *Molecular Biology and Evolution* **1987**, 4, 572.
- [164] U. Wittstock, B. A. Halkier, *Journal of Biological Chemistry* **2000**, 275, 14659.
- [165] C. H. Hansen, U. Wittstock, C. E. Olsen, A. J. Hick, J. A. Pickett, B. A. Halkier, *Journal of Biological Chemistry* **2001**, 276, 11078.
- [166] P. Naur, C. H. Hansen, S. Bak, B. G. Hansen, N. B. Jensen, H. L. Nielsen, B. A. Halkier, *Archives of Biochemistry and Biophysics* **2003**, 409, 235.
- [167] C. Haudenschild, M. Schalk, F. Karp, R. Croteau, *Archives of Biochemistry and Biophysics* **2000**, 379, 127.
- [168] C. M. Berteau, M. Schalk, F. Karp, M. Maffei, R. Croteau, *Archives of Biochemistry and Biophysics* **2001**, 390, 279.
- [169] M. C. Y. Chang, R. A. Eachus, W. Trieu, D. K. Ro, J. D. Keasling, *Nature Chemical Biology* **2007**, 3, 274.
- [170] C. Duport, R. Spagnoli, E. Degryse, D. Pompon, *Nature Biotechnology* **1998**, 16, 186.
- [171] F. M. Szczepara, C. Chandelier, C. Villeret, A. Masurel, S. Bourrot, C. Duport, S. Blanchard, A. Groisillier, E. Testet, P. Costaglioli, G. Cauet, E. Degryse, D. Balbuena, J. Winter, T. Achstetter, R. Spagnoli, D. Pompon, B. Dumas, *Nature Biotechnology* **2003**, 21, 143.
- [172] A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver, *Science* **1996**, 274, 546.
- [173] V.-K. Ton, R. Rao, *American Journal of Physiology and Cell Physiology* **2004**, 287, C580.
- [174] S. Macauley-Patrick, M. L. Fazenda, B. McNeil, L. M. Harvey, *Yeast* **2005**, 22, 249.
- [175] G. P. L. Cereghino, J. L. Cereghino, C. Ilgen, J. M. Cregg, *Current Opinion in Biotechnology* **2002**, 13, 329.
- [176] G. Gellissen, G. Kunze, C. Gaillardin, J. M. Cregg, E. Berardi, M. Veenhuis, I. van der Klei, *FEMS Yeast Research* **2005**, 5, 1079.
- [177] T. Wartmann, U. W. Stephan, I. Bube, E. Böer, M. Melzer, R. Manteuffel, R. Stoltenburg, L. Guengerich, G. Gellissen, G. Kunze, *Yeast* **2002**, 19, 849.
- [178] R. J. Andereg, R. Betz, S. A. Carr, J. W. Crabb, W. Duntze, *Journal of Biological Chemistry* **1988**, 263, 18236.
- [179] M. A. Romanos, C. A. Scorer, J. J. Clare, *Yeast* **1992**, 8, 423.
- [180] A. Hinnen, J. B. Hicks, G. R. Fink, *Proceedings of the National Academy of Sciences of the United States of America* **1978**, 75, 1929.
- [181] T. S. Lopes, J. Klootwijk, A. E. Veenstra, P. C. Vanderaar, H. Vanheerikhuizen, H. A. Raue, R. J. Planta, *Gene* **1989**, 79, 199.
- [182] T. S. Lopes, G. J. A. J. Hakkaart, B. L. Koerts, H. A. Raue, R. J. Planta, *Gene* **1991**, 105, 83.
- [183] A. L. Goldstein, J. H. McCusker, *Yeast* **1999**, 15, 1541.
- [184] S. Ito-Harashima, J. H. McCusker, *Yeast* **2004**, 21, 53.
- [185] G. Jansen, C. L. Wu, B. Schade, D. Y. Thomas, M. Whiteway, *Gene* **2005**, 344, 43.

- [186] I. Sadowski, T. C. Su, J. Parent, *Yeast* **2007**, 24, 447.
- [187] A. W. Murray, J. W. Szostak, *Cell* **1983**, 34, 961.
- [188] L. Clarke, J. Carbon, *Nature* **1980**, 287, 504.
- [189] A. B. Futcher, *Yeast* **1988**, 4, 27.
- [190] T. P. Stjohn, R. W. Davis, *Journal of Molecular Biology* **1981**, 152, 285.
- [191] R. J. Rodriguez, C. Low, C. D. K. Bottema, L. W. Parks, *Biochimica Et Biophysica Acta* **1985**, 837, 336.
- [192] S. J. Smith, J. H. Crowley, L. W. Parks, *Molecular and Cellular Biology* **1996**, 16, 5427.
- [193] F. Mantzouridou, E. Naziri, M. Z. Tsimidou, *Journal of Agricultural and Food Chemistry* **2009**, 57, 6189.
- [194] T. A. Lewis, F. R. Taylor, L. W. Parks, *Journal of Bacteriology* **1985**, 163, 199.
- [195] P. J. Trocha, D. B. Sprinson, *Archives of Biochemistry and Biophysics* **1976**, 174, 45.
- [196] L. Jahnke, H. P. Klein, *Journal of Bacteriology* **1983**, 155, 488.
- [197] A. A. Andreasen, T. J. B. Stier, *Journal of Cellular and Comparative Physiology* **1953**, 41, 23.
- [198] K. E. Kwast, L. C. Lai, N. Menda, D. T. James, S. Aref, P. V. Burke, *Journal of Bacteriology* **2002**, 184, 250.
- [199] V. J. Cid, A. Duran, F. Delrey, M. P. Snyder, C. Nombela, M. Sanchez, *Microbiological Reviews* **1995**, 59, 345.
- [200] T. L. Lewis, G. A. Keesler, G. P. Fenner, L. W. Parks, *Yeast* **1988**, 4, 93.
- [201] J. H. Crowley, F. W. Leak, K. V. Shianna, S. Tove, L. W. Parks, *Journal of Bacteriology* **1998**, 180, 4177.
- [202] B. S. J. Davies, H. S. Wang, J. Rine, *Molecular and Cellular Biology* **2005**, 25, 7375.
- [203] C. R. Woese, L. J. Magrum, G. E. Fox, *Journal of Molecular Evolution* **1978**, 11, 245.
- [204] M. Hanford, T. L. Peeples, *Applied Biochemistry and Biotechnology* **2002**, 97, 45.
- [205] a) M. Derosa, A. Gambacorta, A. Gliozzi, *Microbiological Reviews* **1986**, 50, 70; b) M. DeRosa, *Thin Solid Films* **1996**, 285, 13.
- [206] J. S. Lolkema, G. Speelmans, W. N. Konings, *Biochimica Et Biophysica Acta-Bioenergetics* **1994**, 1187, 211.
- [207] T. Benvegnu, L. Lemiegre, S. Cammas-Marion, *European Journal of Organic Chemistry* **2008**, 4725.
- [208] B. A. Kellogg, C. D. Poulter, *Current Opinion in Chemical Biology* **1997**, 1, 570.
- [209] A. J. Chen, C. D. Poulter, *Journal of Biological Chemistry* **1993**, 268, 11002.
- [210] a) S. Ohnuma, M. Suzuki, T. Nishino, *Journal of Biological Chemistry* **1994**, 269, 14792; b) A. Tachibana, *FEBS Letters* **1994**, 341, 291.
- [211] E. I. Fukusaki, T. Nishikawa, K. Kato, A. Shinmyo, H. Hemmi, T. Nishino, A. Kobayashi, *Journal of Bioscience and Bioengineering* **2003**, 95, 283.
- [212] L. M. Chen, K. Brugger, M. Skovgaard, P. Redder, Q. X. She, E. Torarinsson, B. Greve, M. Awayez, A. Zibat, H. P. Klenk, R. A. Garrett, *Journal of Bacteriology* **2005**, 187, 4992.
- [213] S. K. Eszterhas, E. E. Bouhassira, D. I. K. Martin, S. Fiering, *Molecular and Cellular Biology* **2002**, 22, 469.
- [214] V. Pelechano, J. Garcia-Martinez, J. E. Perez-Ortin, *Yeast* **2006**, 23, 689.
- [215] B. Dujon, *Trends in Genetics* **1996**, 12, 263.

- [216] C. Springer, O. Valerius, A. Strittmatter, G. H. Braus, *Journal of Biological Chemistry* **1997**, 272, 26318.
- [217] O. Valerius, C. Brendel, K. Düvel, G. H. Braus, *Journal of Biological Chemistry* **2002**, 277, 21440.
- [218] a) I. H. Greger, N. J. Proudfoot, *Embo Journal* **1998**, 17, 4771; b) I. H. Greger, A. Aranda, N. Proudfoot, *Proceedings of the National Academy of Sciences of the United States of America* **2000**, 97, 8415.
- [219] S. Puig, J. E. Perez-Ortin, E. Matallana, *Current Microbiology* **1999**, 39, 369.
- [220] E. M. Prescott, N. J. Proudfoot, *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, 8796.
- [221] D. Mumberg, R. Müller, M. Funk, *Gene* **1995**, 156, 119.
- [222] K. Tokuhira, M. Muramatsu, C. Ohto, T. Kawaguchi, S. Obata, N. Muramoto, M. Hirai, H. Takahashi, A. Kondo, E. Sakuradani, S. Shimizu, *Applied and Environmental Microbiology* **2009**, 75, 5536.
- [223] K. R. Oldenburg, K. T. Vo, S. Michaelis, C. Paddon, *Nucleic Acids Research* **1997**, 25, 451.
- [224] C. Schaerer-Brodbeck, A. Barberis, *Biotechniques* **2004**, 37, 202.
- [225] M. Itaya, K. Fujita, A. Kuroki, K. Tsuge, *Nature Methods* **2008**, 5, 41.
- [226] D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nature Methods* **2009**, 6, 343.
- [227] D. Cohen, I. Chumakov, J. Weissenbach, *Nature* **1993**, 366, 698.
- [228] C. Bellanné-Chantelot, B. Lacroix, P. Ougen, A. Billault, S. Beauvils, S. Bertrand, I. Georges, F. Glibert, I. Gros, G. Lucotte, L. Susini, J.-J. Codani, P. Gesnouin, S. Pook, G. Vaysseix, J. Lu-Kuo, T. Ried, D. Ward, I. Chumakov, D. Le Paslier, E. Barillot, D. Cohen, *Cell* **1992**, 70, 1059.
- [229] H. Takeda, H. Yamakuchi, N. Ihara, K. Hara, T. Watanabe, Y. Sugimoto, T. Oshiro, H. Kishine, Y. Kano, K. Kohno, *Animal Genetics* **1998**, 29, 216.
- [230] C. RogelGaillard, N. Bourgeaux, J. C. Save, C. Renard, P. Coullin, P. Pinton, M. Yerle, M. Vaiman, P. Chardon, *Mammalian Genome* **1997**, 8, 186.
- [231] M. Naesby, S. V. S. Nielsen, C. A. F. Nielsen, T. Green, T. O. Tange, E. Simon, P. Knechtle, A. Hansson, M. S. Schwab, O. Titiz, C. Folly, R. E. Archila, M. Maver, S. V. Fiet, T. Boussemghoune, M. Janes, A. S. S. Kumar, S. P. Sonkar, P. P. Mitra, V. A. K. Benjamin, N. Korrapati, I. Suman, E. H. Hansen, T. Thybo, N. Goldsmith, A. S. Sorensen, *Microbial Cell Factories* **2009**, 8.
- [232] M. Cocchia, S.-J. Kouprina, V. Larionov, D. Schlessinger, R. Nagaraja, *Nucleic Acids Research* **2000**, 28.
- [233] R. Verwaal, J. Wang, J. P. Meijnen, H. Visser, G. Sandmann, J. A. van den Berg, A. J. J. van Ooyen, *Applied and Environmental Microbiology* **2007**, 73, 4342.
- [234] Y. Jiang, P. Proteau, D. Poulter, S. Ferronovick, *Journal of Biological Chemistry* **1995**, 270, 21793.
- [235] K. Ukibe, K. Hashida, N. Yoshida, H. Takagi, *Applied and Environmental Microbiology* **2009**, 75, 7205.
- [236] J. Maury, M. A. Asadollahi, K. Moller, M. Schalk, A. Clark, L. R. Formenti, J. Nielsen, *FEBS Letters* **2008**, 582, 4032.

- [237] A. Das, S. H. Yoon, S. H. Lee, J. Y. Kim, D. K. Oh, S. W. Kim, *Applied Microbiology and Biotechnology* **2007**, 77, 505.
- [238] S. H. Yoon, Y. M. Lee, J. E. Kim, S. H. Lee, J. H. Lee, J. Y. Kim, K. H. Jung, Y. C. Shin, J. D. Keasling, S. W. Kim, *Biotechnology and Bioengineering* **2006**, 94, 1025.
- [239] Y. Shiba, E. M. Paradise, J. Kirby, D. K. Ro, J. D. Keasling, *Metabolic Engineering* **2007**, 9, 160.
- [240] D. K. Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. M. Ndungu, K. A. Ho, R. A. Eachus, T. S. Ham, J. Kirby, M. C. Y. Chang, S. T. Withers, Y. Shiba, R. Sarpong, J. D. Keasling, *Nature* **2006**, 440, 940.
- [241] M. A. Asadollahi, J. Maury, M. Schalk, A. Clark, J. Nielsen, *Biotechnology and Bioengineering* **2010**, 106, 86.
- [242] T. Polakowski, U. Stahl, C. Lang, *Applied Microbiology and Biotechnology* **1998**, 49, 66.
- [243] G. Laskaris, R. van der Heijden, R. Verpoorte, *Plant Science* **2000**, 153, 97.
- [244] A. Grote, K. Hiller, M. Scheer, R. Munch, B. Nortemann, D. C. Hempel, D. Jahn, *Nucleic Acids Research* **2005**, 33, W526.
- [245] S. Takahashi, Y. Yeo, B. T. Greenhagen, T. McMullin, L. Song, J. Maurina-Brunker, R. Rosson, J. P. Noel, J. Chappell, *Biotechnology and Bioengineering* **2007**, 97, 170.
- [246] M. A. Asadollahi, J. Maury, K. Moller, K. F. Nielsen, M. Schalk, A. Clark, J. Nielsen, *Biotechnology and Bioengineering* **2008**, 99, 666.
- [247] E. Leonard, P. K. Ajikumar, K. Thayer, W.-H. Xiao, J. D. Mo, B. Tidor, G. Stephanopoulos, K. L. J. Prather, *Proceedings of the National Academy of Sciences* **2010**, 107, 13654.
- [248] A. Faulkner, X. M. Chen, J. Rush, B. Horazdovsky, C. J. Waechter, G. M. Carman, P. C. Sternweis, *Journal of Biological Chemistry* **1999**, 274, 14831.
- [249] W. I. Wu, Y. S. Liu, B. Riedel, J. B. Wissing, A. S. Fischl, G. M. Carman, *Journal of Biological Chemistry* **1996**, 271, 1868.
- [250] A. Tachibana, T. Tanaka, M. Taniguchi, S. Oi, *FEBS Letters* **1996**, 379, 43.
- [251] D. W. Christianson, *Chemical Reviews* **2006**, 106, 3412.
- [252] P. R. Wilderman, R. J. Peters, *Journal of the American Chemical Society* **2007**, 129, 15736.
- [253] M. M. Xu, P. R. Wilderman, R. J. Peters, *Proceedings of the National Academy of Sciences of the United States of America* **2007**, 104, 7397.
- [254] J. E. Mertz, R. W. Davis, *Proceedings of the National Academy of Sciences of the United States of America* **1972**, 69, 3370.
- [255] G. Q. Chen, I. Choi, B. Ramachandran, J. E. Gouaux, *Journal of the American Chemical Society* **1994**, 116, 8799.
- [256] W. P. C. Stemmer, A. Cramer, K. D. Ha, T. M. Brennan, H. L. Heyneker, *Gene* **1995**, 164, 49.
- [257] M. Itaya, K. Fujita, M. Ikeuchi, M. Koizumi, K. Tsuge, *Journal of Biochemistry* **2003**, 134, 513.
- [258] Z. Y. Shao, H. Zhao, H. M. Zhao, *Nucleic Acids Research* **2009**, 37.
- [259] M. Z. Li, S. J. Elledge, *Nature Methods* **2007**, 4, 251.
- [260] J. Bitinaite, M. Rubino, K. H. Varma, I. Schildkraut, R. Vaisvila, R. Vaiskunaite, *Nucleic Acids Research* **2007**, gkm041.
- [261] B. G. Zhu, G. F. Cai, E. O. Hall, G. J. Freeman, *Biotechniques* **2007**, 43, 356.



- [262] D. G. Gibson, G. A. Benders, C. Andrews-Pfannkoch, E. A. Denisova, H. Baden-Tillson, J. Zaveri, T. B. Stockwell, A. Brownley, D. W. Thomas, M. A. Algire, C. Merryman, L. Young, V. N. Noskov, J. I. Glass, J. C. Venter, C. A. Hutchison, H. O. Smith, *Science* **2008**, *319*, 1215.
- [263] Y. Aylon, M. Kupiec, *Molecular and Cellular Biology* **2003**, *23*, 6585.
- [264] J. N. Strathern, A. J. S. Klar, J. B. Hicks, J. A. Abraham, J. M. Ivy, K. A. Nasmyth, C. McGill, *Cell* **1982**, *31*, 183.
- [265] D. L. Hu, M. Crist, X. Q. Duan, F. S. Gimble, *Biochemistry* **1999**, *38*, 12621.
- [266] A. Plessis, A. Perrin, J. E. Haber, B. Dujon, *Genetics* **1992**, *130*, 451.
- [267] S. N. Guzder, Y. Habraken, P. Sung, L. Prakash, S. Prakash, *Journal of Biological Chemistry* **1995**, *270*, 12973.
- [268] M. Sabourin, J. L. Nitiss, K. C. Nitiss, K. Tatebayashi, H. Ikeda, N. Osherooff, *Nucleic Acids Research* **2003**, *31*, 4373.
- [269] B. O. Krogh, L. S. Symington, *Annual Review of Genetics* **2004**, *38*, 233.
- [270] A. Barzel, M. Kupiec, *Nature Reviews Genetics* **2008**, *9*, 27.
- [271] T. Allers, M. Lichten, *Molecular Cell* **2001**, *8*, 225.
- [272] J. W. Szostak, T. L. Orr-Weaver, R. J. Rothstein, F. W. Stahl, *Cell* **1983**, *33*, 25.
- [273] J. S. Filipp, P. Sung, H. Klein, *Annual Review of Biochemistry* **2008**, *77*, 229.
- [274] N. Nassif, J. Penney, S. Pal, W. R. Engels, G. B. Gloor, *Molecular and Cellular Biology* **1994**, *14*, 1613.
- [275] L. S. Symington, *Microbiological and Molecular Biology Reviews* **2002**, *66*, 630.
- [276] J. M. Daley, P. L. Palmbo, D. Wu, T. E. Wilson, *Annual Review of Genetics* **2005**, *39*, 431.
- [277] M. McVey, S. E. Lee, *Trends in Genetics* **2008**, *24*, 529.
- [278] M. T. Madigan, J. Martinko, J. Parker (Ed.: W. Goebel) *Mikrobiologie* **2001**, Spektrum Akademischer Verlag GmbH, Heidelberg.
- [279] M. Szostková, D. Horáková, *Bioelectrochemistry and Bioenergetics* **1998**, *47*, 319.
- [280] M. Walfridsson, J. Hallborn, M. Penttila, S. Keranen, B. Hahn-Hagerdal, *Applied and Environmental Microbiology* **1995**, *61*, 4184.
- [281] P. V. Lograsso, D. A. Soltis, B. R. Boettcher, *Archives of Biochemistry and Biophysics* **1993**, *307*, 193.
- [282] S. Picaud, L. Olofsson, M. Brodelius, P. E. Brodelius, *Archives of Biochemistry and Biophysics* **2005**, *436*, 215.
- [283] S. M. Jennings, Y. H. Tsay, T. M. Fisch, G. W. Robinson, *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 6038.
- [284] T. Satoh, M. Horie, H. Watanabe, Y. Tsuchiya, T. Kamei, *Biological & Pharmaceutical Bulletin* **1993**, *16*, 349.
- [285] M. Veen, U. Stahl, C. Lang, *FEMS Yeast Research* **2003**, *4*, 87.
- [286] H. Alper, C. Fischer, E. Nevoigt, G. Stephanopoulos, *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 12678.
- [287] B. F. Pfleger, D. J. Pitera, C. D. Smolke, J. D. Keasling, *Nature Biotechnology* **2006**, *24*, 1027.

## VII Anhang

### Abkürzungsverzeichnis

2µm	2 micron
<i>A. thaliana</i>	<i>Arabidopsis thaliana</i>
Ack	Acetatkinase
ADP	Adenosindiphosphat
AMP	Adenosinmonophosphat
AmpR	Ampicillinresistenzgen
AP	Alkalische Phosphatase
ARS	<i>autonomous replication site</i> (Hefereplikationsursprung)
AS	Aminosäure
ATCC	<i>American Type Culture Collection</i>
ATP	Adenosintriphosphat
<i>attB/attP/attL/attR</i>	<i>attachment-sites</i> des Gateway-Systems
BOC	<i>tert</i> -Butyloxycarbonyl
bp	Basenpaar
BP-Reaktion	Rekombination zwischen <i>attB</i> - und <i>attP</i> -Sequenzen
BSA	<i>Bovine serum albumin</i> (Rinderserum)
bzw.	Beziehungsweise
C13T	C13- <i>O</i> -Phenylpropanoyltransferase
ccdB	<i>cell cycle division</i> (Gen für Gyrasehemmer)
cDNS	<i>complementary</i> Desoxyribonukleinsäure
CEN	Zentromersequenz

CmR	Chloramphenicolresistenz
CO <sub>2</sub>	Kohlenstoffdioxid
CoA	Coenzym A
C-Quelle	Kohlenstoffquelle
cycltt	Terminatorregion des Cytochrom 1-Gens
ddH <sub>2</sub> O	doppelt destilliertes Wasser
ddNTP	Didesoxyribonukleinsäure-Triphosphat
<i>de novo</i>	vom Anfang
DNS	Desoxyribonukleinsäure
dNTP	Desoxyribonukleinsäure-Trisphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
ER	Endoplasmatisches Reticulum
ERG9	Squalensynthase
<i>et al.</i>	<i>et alii</i> (und andere)
EUROSCARF	<i>European S. cerevisiae archive for functional analysis</i>
FDA	<i>Food and Drug Administration</i>
for	<i>forward</i> (vorwärts)
gap1p	Promotor des Glycerinaldehyd-3-P-dehydrogenasegens
GC	Gaschromatographie
gcn4tt	Terminator des Transkriptionsaktivatorgens
GGPPS	Geranylgeranyldiphosphatsynthase
GGPPSco	für die Expression in <i>S. cerevisiae</i> optimierte GGPPS
GGPPS-Sa	GGPP aus <i>Sulfolobus acidocaldarius</i>



---

GGPPS-Tc	GGPPS aus <i>T. chinensis</i>
gdp	Promotor des Glukose-6-phosphat-dehydrogenasegens
GRAS	<i>Generally regarded as safe</i>
GW	Gateway
HIS	Histidin
HMGR	HMG-CoA-Reduktase
<i>in vitro</i>	außerhalb eines lebenden Organismus
<i>in vivo</i>	im lebenden Organismus
Inv	invers/invertiert
IS	Interner Standard
kb	Kilobasenpaare
KmR	Kanamycinresistenzgen
LacZ	$\beta$ -Galaktosidase zur Blau-weiß-Selektion
LEU	Leucin
LR-Reaktion	Rekombination zwischen <i>attL</i> - und <i>attR</i> -Sequenzen
MAT	<i>Mating type</i>
MCS	<i>Multiple cloning site</i>
mRNS	<i>messenger RNS</i>
MS	Massenspektrometrie
n.d.	nicht detektierbar
NADPH	Nikotinsäureamid-Adenin-Dinukleotid-Phosphat
NEB	New England Biolabs
N-terminal	Aminosäureende eines Proteins /Peptids

---

NY	New York
OAA	Oxalacetat
ORF	<i>open reading frame</i>
PAA	Polyacrylamid
PAGE	Polyacrylamidgelelektrophorese
PCI	Phenol-Chlorophorm-Isoamylalkohol (25:24:1)
Pck	PEP-Carboxykinase
PCR	<i>Polymerase chain reaction</i> (Polymerasekettenreaktion)
PEP	Phosphoenolpyruvat
pgkp	Promotor des Phosphoglyceratkinasgens
pho5tt	Terminator des Sauren Phosphatasgens
Ppc	PEP-Carboxylase
Pps	PEP-Synthase
Pta	Phosphotransacetylase
Pyk	Pyruvatkinase
rev	<i>reverse</i> (rückwärts)
RNS	Ribonukleinsäure
RP	<i>reversed phase</i>
rRNS	ribosomale RNS
RU	<i>relative unit</i> (relative Einheit)
<i>S. acidocaldarius</i>	<i>Sulfolobus acidocaldarius</i>
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	<i>Sodium dodecyl sulfate</i>

SOE-PCR	<i>Splicing by overlap extension PCR</i>
spp.	Spezies
SQ	Squalen
<i>T. baccata</i>	<i>Taxus baccata</i>
<i>T. canadensis</i>	<i>Taxus canadensis</i>
<i>T. chinensis</i>	<i>Taxus chinensis</i>
<i>T. cuspidata</i>	<i>Taxus cuspidata</i>
T10Ac	10-Deacetylbaecatin III-10- <i>O</i> -acetyltransferase
T10H	Taxan-10 $\beta$ -hydroxylase
T13H	Taxan-13 $\alpha$ -hydroxylase
T14H	Taxan-14 $\beta$ -hydroxylase
T2Ac	2 $\alpha$ - <i>O</i> -Benzoyltransferase
T2H	Taxoid-2 $\alpha$ -hydroxylase
T5Ac	Taxadien-5 $\alpha$ - <i>O</i> -acetyltransferase
T5H	Taxadien-5 $\alpha$ -hydroxylase
T7H	Taxoid-7 $\beta$ -hydroxylase
TCA	Trichloressigsäure
TCS	Tricarbonsäure
TD	Taxadien
TDS	Taxadiensynthase
TDSco	für die Expression in <i>S. cerevisiae</i> optimierte TDS
TDSco-80	um 80 AS verkürzte TDSco
TDS-Tc	Taxadiensynthase aus <i>Taxus chinensis</i>

---

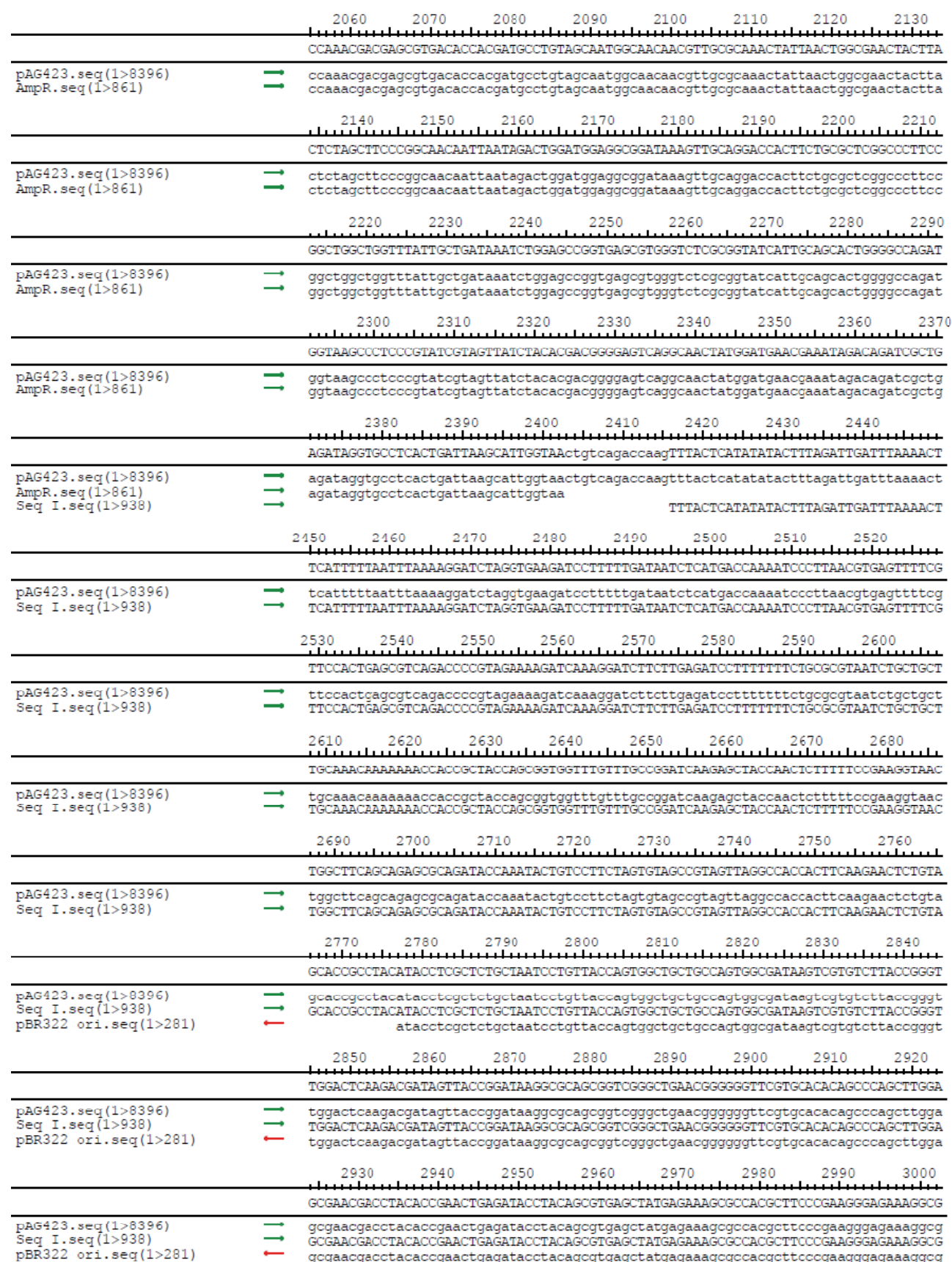
tefp	Promotor des Transkriptionselongationsfaktors
TEL	Telomersequenzen
tHMGR	um die regulatorische Domäne verkürzte HMGR
TI	Transkriptionale Interferenz
T <sub>m</sub>	Schmelzpunkt
TRP	Tryptophan
UPC2	Transkriptionsfaktor <i>uptake control</i>
UPC2.1	punktmutierte Version von UPC2
URA	Uracil
USA	<i>United States of America</i>
WHO	<i>World Health Organisation</i>
Wt	Wildtyp
YAC	<i>Yeast artificial chromosome</i> (künstliches Hefechromosom)

Genomic alignment of pAG423 and 2micron sequences to the E. coli chromosome. The figure displays 15 rows of sequence alignments, each showing a reference sequence (top) and a query sequence (bottom). The reference sequence is the E. coli chromosome, and the query sequence is either pAG423 (1>8396) or 2micron (1>1161). The alignment is indicated by green arrows. The sequences are aligned to the E. coli chromosome, which is represented by a horizontal line with tick marks indicating positions. The positions are labeled at the top and bottom of the alignment.

Row	Reference Sequence (E. coli)	Query Sequence
1	gacgaaagggcctcgtgatacgcctatttttatagggttaatgtcatgataataatgggtttccttagatgatccaatatca	pAG423.seq (1>8396)
2	aaggaaatgatagcattgaaggatgagactaatccaattgaggagtgccagcatatagaacagctaaagggtagtgctg	pAG423.seq (1>8396)
3	aaggaagcatacgatacccgcatggaatgggataaatatcacaggaggtactagactacctttcatcctacataaatag	pAG423.seq (1>8396)
4	acgcatataagtagcatttaagcataaaacagcactatgcccgttcttctcatgtatatatatatacaggcaaacagca	pAG423.seq (1>8396) 2micron.seq (1>1161)
5	gatataggtgacgacgtgaacagtgagctgtatgtgcccagctcgcgttgcatttttcggaagcgctcgttttcggaaacg	pAG423.seq (1>8396) 2micron.seq (1>1161)
6	ctttgaagttcctattccgaagttcctattctctagaaagtataaggaacttcagagcgcttttgaaaacccaaagcgct	pAG423.seq (1>8396) 2micron.seq (1>1161)
7	ctgaagacgcacttttcaaaaaacccaaacgcacggactgtaacgagctactaaaattatgogaatcacgcgttcocaca	pAG423.seq (1>8396) 2micron.seq (1>1161)
8	aacattgctcaaaagtatctctttgctatatatctctgtgctatatccctatataacctacccatccacctttcgctcc	pAG423.seq (1>8396) 2micron.seq (1>1161)
9	ttgaacttgcatctaaactcgacctctacattttttatggttatctctagttactcttttagacaaaaaattgtagt	pAG423.seq (1>8396) 2micron.seq (1>1161)
10	aagaactattcatagagtgaaatcgaaaacaatcgaaaatgtaaacatttctctatagtagtatatagagacaaaatag	pAG423.seq (1>8396) 2micron.seq (1>1161)
11	aagaaacoggttcataattttctgaccaatgaagaatcatcaacgctatcactttctgttcacaaagtatgcgcaatcca	pAG423.seq (1>8396) 2micron.seq (1>1161)
12	catcggtatagaatataatcggggatgcctttatcttgaaaaaatgcaccgcagcttcgctagtaaatcagtaaacgcg	pAG423.seq (1>8396) 2micron.seq (1>1161)
13	ggaagtggagtcaggctttttttatggaagagaaaatagacaccaaagtagccttcttcttaaccttaacggacctacag	pAG423.seq (1>8396) 2micron.seq (1>1161)

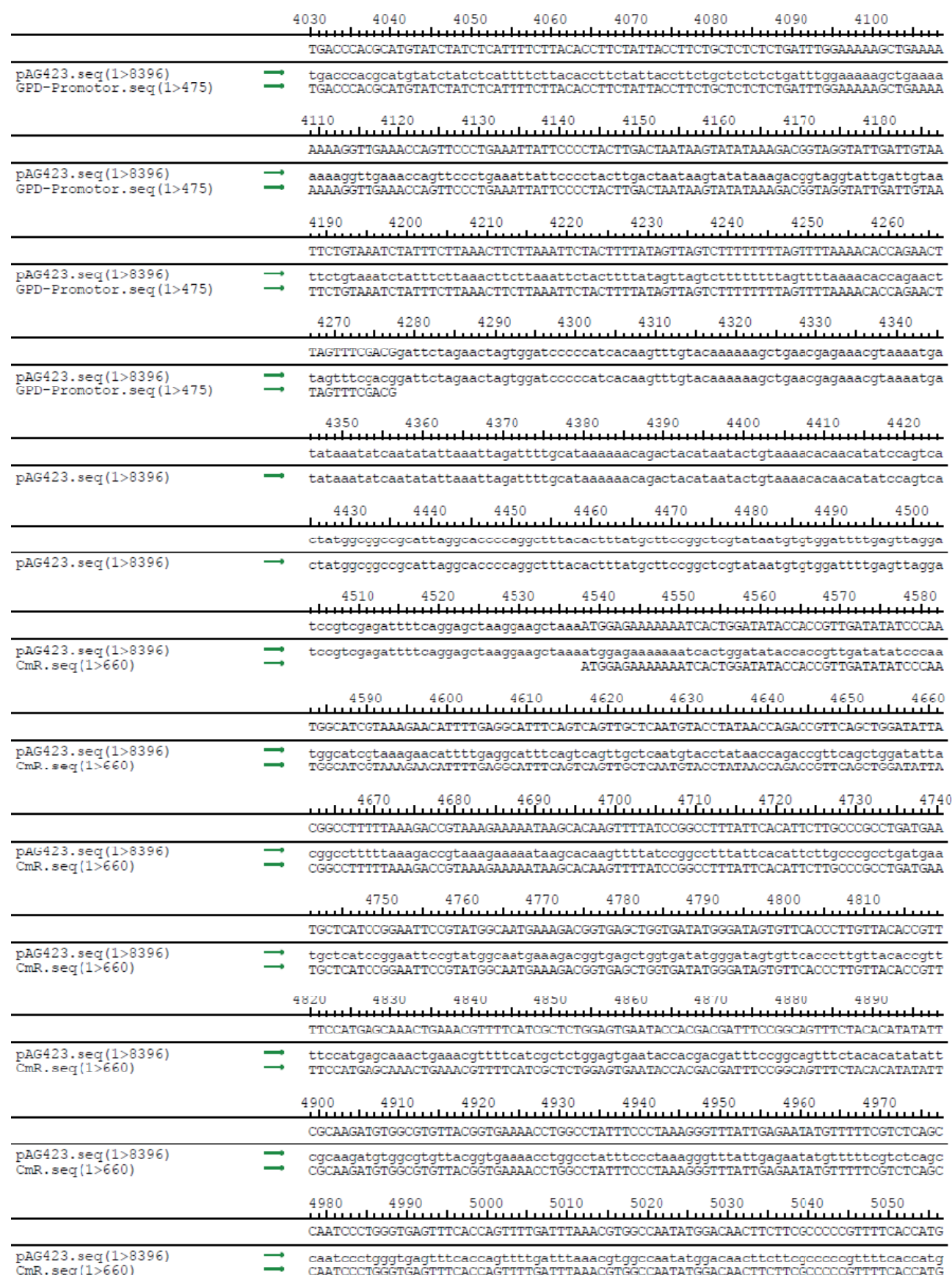
			1030	1040	1050	1060	1070	1080	1090	1100
			TGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAGTAATCTAAGATGCTTTGTTAGAAAAA							
pAG423.seq(1>8396)	→		tgcaaaaagttatcaagagactgcattatagagcgccacaaaggagaaaaaaagtaaatctaagatgctttgtagaaaaa							
2micron.seq(1>1161)	→		tgcaaaaagttatcaagagactgcattatagagcgccacaaaggagaaaaaaagtaaatctaagatgctttgtagaaaaa							
			1110	1120	1130	1140	1150	1160	1170	1180
			TAGCGCTCTCGGGATGCATTTTGTAGAACAAAAAGAAGTATAGATTCTTTGTTGGTAAATAGCGCTCTCGCGTTGC							
pAG423.seq(1>8396)	→		tagcgctctcgggatgcatttttgtagaacaaaaagaagtatagattctttgttggttaaatagcgctctcgcgttgc							
2micron.seq(1>1161)	→		tagcgctctcgggatgcatttttgtagaacaaaaagaagtatagattctttgttggttaaatagcgctctcgcgttgc							
			1190	1200	1210	1220	1230	1240	1250	1260
			ATTTCGTGTTCTGTAATAATGCAGCTCAGATTCTTTGTTTGAATAATAGCGCTCTCGCGTTGCATTTTGTGTTTACAAA							
pAG423.seq(1>8396)	→		atttctgttctgtaaaaatgcagctcagattctttgtttgaaaaattagcgctctcgcggttcatttttgttttacaaa							
2micron.seq(1>1161)	→		atttctgttctgtaaaaatgcagctcagattctttgtttgaaaaattagcgctctcgcggttcatttttgttttacaaa							
			1270	1280	1290	1300	1310	1320	1330	1340
			AATGAAGCACAGATTCTTCGTTGGTAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAATAATGCAGCTCAGATTCT							
pAG423.seq(1>8396)	→		aatgaagcacagattcttcggttgtaaaaatagcgctttcgcggttcattttctgttctgtaaaaatgcagctcagattct							
2micron.seq(1>1161)	→		aatgaagcacagattcttcggttgtaaaaatagcgctttcgcggttcattttctgttctgtaaaaatgcagctcagattct							
			1350	1360	1370	1380	1390	1400	1410	1420
			TTGTTTGAATAATAGCGCTCTCGCGTTGCATTTTGTCTACAAAATGAAGCACAGATGCTTCGTTcaggtggcactt							
pAG423.seq(1>8396)	→		ttgtttgaaaaattagcgctctcgcggttcattttgttctacaaaatgaagcacagatgcttcggtcaggtggcactt							
2micron.seq(1>1161)	→		ttgtttgaaaaattagcgctctcgcggttcattttgttctacaaaatgaagcacagatgcttcggtt							
			1430	1440	1450	1460	1470	1480	1490	1500
			ttcggggaaatgtgcgoggaacccctatttgtttatttttctaaatacattcaaatatgtatcgcgtcatgagacaata							
pAG423.seq(1>8396)	→		ttcggggaaatgtgcgoggaacccctatttgtttatttttctaaatacattcaaatatgtatcgcgtcatgagacaata							
			1510	1520	1530	1540	1550	1560	1570	1580
			accctgataaatgcttcaataaattgaaaaaggaagagtATGAGTATTCAACATTTCOGTGTGCGCCCTATTCCOCTTT							
pAG423.seq(1>8396)	→		accctgataaatgcttcaataaattgaaaaaggaagagtatgagtattcaacatttcogtgtgcgccctattcccttt							
AmpR.seq(1>861)	→		atgagtattcaacatttcogtgtgcgccctattcccttt							
			1590	1600	1610	1620	1630	1640	1650	
			TTTGCGGCATTTTGCCCTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTG							
pAG423.seq(1>8396)	→		tttgoggcatttttgcccttcctgtttttgtctaccccagaaacgctggtgaaagtaaagatgctgaagatcagttgggtg							
AmpR.seq(1>861)	→		tttgoggcatttttgcccttcctgtttttgtctaccccagaaacgctggtgaaagtaaagatgctgaagatcagttgggtg							
			1660	1670	1680	1690	1700	1710	1720	1730
			CACGAGTGGGTACATCGAAGTCTCAACAGCGGTAGATCCTTGAGAGTTTTCGCCCGAAGAACGTTTCCCAAT							
pAG423.seq(1>8396)	→		cacgagtggtgttacatogaactggatctcaacagcggtaagatcccttgagagttttgcgcccgaaagaacgttttccaat							
AmpR.seq(1>861)	→		cacgagtggtgttacatogaactggatctcaacagcggtaagatcccttgagagttttgcgcccgaaagaacgttttccaat							
			1740	1750	1760	1770	1780	1790	1800	1810
			GATGAGCACTTTTAAAGTTCTGCTATGTGGCGGGTATTATCCCGTATTGAGCGCCGGGCAAGAGCAACTCGGTGCGCGC							
pAG423.seq(1>8396)	→		gatgagcacttttaaaagtctgtatgtggcggttattatccogtattgagcgcggggaagagcaactcggtcgcgc							
AmpR.seq(1>861)	→		gatgagcacttttaaaagtctgtatgtggcggttattatccogtattgagcgcggggaagagcaactcggtcgcgc							
			1820	1830	1840	1850	1860	1870	1880	1890
			ATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAG							
pAG423.seq(1>8396)	→		atacactattctcagaatgacttgggttgagtactcaccagtcacagaaaagcatcttaacggatggcatgacagtaagag							
AmpR.seq(1>861)	→		atacactattctcagaatgacttgggttgagtactcaccagtcacagaaaagcatcttaacggatggcatgacagtaagag							
			1900	1910	1920	1930	1940	1950	1960	1970
			AATTATGCACTGCTGCATACCATGAGTGATAACACTGCGGCCAAGTACTTCTGACAAAGCATCGGAGGACCGAAGGA							
pAG423.seq(1>8396)	→		aattatgcagtgctgccataaaccatgagtgataaacactcgggccaacttactctgacaaacgatcgggaggaacgaagga							
AmpR.seq(1>861)	→		aattatgcagtgctgccataaaccatgagtgataaacactcgggccaacttactctgacaaacgatcgggaggaacgaagga							
			1980	1990	2000	2010	2020	2030	2040	2050
			GCTAACCGCTTTTTTGCACAACTGCGGGATCATGTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATA							
pAG423.seq(1>8396)	→		gctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaacccggagctgaatgaagccata							
AmpR.seq(1>861)	→		gctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaacccggagctgaatgaagccata							







			3010 3020 3030 3040 3050 3060 3070 3080
			GACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGGAAACGGCTGGTATCTTT
pAG423.seq(1>8396)	→		gacaggtatccggtaagcggcagggtcggaacaggagagcgacaggggagcttccaggggaaacgctggtatcttt
Seq I.seq(1>938)	→		GACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTCCAGGGGGAAACGGCTGGTATCTTT
pBR322 ori.seq(1>281)	→		gacaggtatccggtaagcggcagggtcggaacaggagagcgacaggggagcttc
			3090 3100 3110 3120 3130 3140 3150 3160
			ATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAA
pAG423.seq(1>8396)	→		atagtcctgtcgggttttcgccacctctgacttgagcgtcgatTTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAA
Seq I.seq(1>938)	→		ATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGCGGAGCCTATGGAA
			3170 3180 3190 3200 3210 3220 3230
			AAACGCCAGCAACGCGGCTTTTACGGTTCCTGGCCTTTTGTGCTGCTCATGTCTTTCTCTCGCTTATCC
pAG423.seq(1>8396)	→		aaacgccagcaacgcggtcttttacggttcctggccttttgcctggccttttgcctcacatgttcttctcgtggtatcc
Seq I.seq(1>938)	→		AAACGCCAGCAACGCGGCTTTTACGGTTCCTGGCCTTTTGTGCTGCTCATGTCTTTCTCTCGCTTATCC
			3240 3250 3260 3270 3280 3290 3300 3310
			CCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGCGCAGCG
pAG423.seq(1>8396)	→		cctgattctgtggataacogtattacogcctttgagtgagctgataccogctcgccgagccgaacgacgagcgagcg
Seq I.seq(1>938)	→		CCTGATTCTGTGGATAACCGTATTACCGCTTTGAGTGAGCTGATACCGCTCGCCGACGCCGAACGACCGCGCAGCG
			3320 3330 3340 3350 3360 3370 3380 3390
			AGTCAGTGAGCGGAGGAAGCGGAAGACGCCAATAAcgcaaacgcctctccccgcgctgtggcgattcattaatgcag
pAG423.seq(1>8396)	→		agtcagtgagcgaggaagcggaagagcgcccaatacgcgaacgcctctccccgcgctgtggcgattcattaatgcag
Seq I.seq(1>938)	→		AGTCAGTGAGCGGAGGAAGCGGAAGACGCCAATA
			3400 3410 3420 3430 3440 3450 3460 3470
			ctggcacgacaggtttcccgactggaagcgggcagtgagcgcaacgcaattaatgtgagttacctcactcattagcca
pAG423.seq(1>8396)	→		ctggcacgacaggtttcccgactggaagcgggcagtgagcgcaacgcaattaatgtgagttacctcactcattagcca
			3480 3490 3500 3510 3520 3530 3540 3550
			ccccaggctttacactttatgcttccggctcctatgttgtgtggaattgtgagcgagataacaatttcacacaggaaaca
pAG423.seq(1>8396)	→		ccccaggctttacactttatgcttccggctcctatgttgtgtggaattgtgagcgagataacaatttcacacaggaaaca
			3560 3570 3580 3590 3600 3610 3620 3630
			gctatgaccatgattacgccaagcgcgcaattaacccctcactaaaggggaacaaaagctggagctcttogagtttatcat
pAG423.seq(1>8396)	→		gctatgaccatgattacgccaagcgcgcaattaacccctcactaaaggggaacaaaagctggagctcttogagtttatcat
			3640 3650 3660 3670 3680 3690 3700 3710
			tatcaatactcgccatttcaaagaatacgtaaataattaatagtagtgatTTTCTCTAactttatttagtcaaaaaatta
pAG423.seq(1>8396)	→		tatcaatactcgccatttcaaagaatacgtaaataattaatagtagtgatTTTCTCTAactttatttagtcaaaaaatta
			3720 3730 3740 3750 3760 3770 3780 3790
			gccttttaattctgctgtaaccogtacatgccaaaatagggggcggttacacagaatatataacacogaaggtgcct
pAG423.seq(1>8396)	→		gccttttaattctgctgtaaccogtacatgccaaaatagggggcggttacacagaatatataacacogaaggtgcct
			3800 3810 3820 3830 3840 3850 3860 3870
			gggtgaacagGTTTATTCCTGGCATCCACTAATATAATGGAGCCCGCTTTTAAAGCTGGCATCCAGAAAAAAGAA
pAG423.seq(1>8396)	→		gggtgaacaggtttattcctggtcatccactaaatataatggagcccgctttttaagctggcatccagaaaaaaagaa
GPD-Promotor.seq(1>475)	→		GTTTATTCCTGGCATCCACTAATATAATGGAGCCCGCTTTTAAAGCTGGCATCCAGAAAAAAGAA
			3880 3890 3900 3910 3920 3930 3940 3950
			TCCCAGCACCAAAATATTGTTTCTTCAACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGG
pAG423.seq(1>8396)	→		tcccagcaccaaaatattgTTTTCTTCAACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGG
GPD-Promotor.seq(1>475)	→		TCCCAGCACCAAAATATTGTTTCTTCAACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGG
			3960 3970 3980 3990 4000 4010 4020
			GGCACAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCGCTGGAGTAAATGATGACACAAGGCAAT
pAG423.seq(1>8396)	→		ggcacaacaggcaaaaaacgggcacaacctcaatggagtgatgcaacctgCGCTGGAGTAAATGATGACACAAGGCAAT
GPD-Promotor.seq(1>475)	→		GGCACAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCGCTGGAGTAAATGATGACACAAGGCAAT



5060 5070 5080 5090 5100 5110 5120 5130

GGCAAATATTATACGCAAGGCGACAGGTGCTGATGCCGCTGGGOGATTACAGGTTTCATCATGCCGCTCTGTGATGGCTTCC

pAG423.seq(1>8396)  
CmR.seq(1>660)

5140 5150 5160 5170 5180 5190 5200 5210

ATGTGCGCAGAATGCTTAATGAATTACACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAacgcgcgcgtggtatccgg

pAG423.seq(1>8396)  
CmR.seq(1>660)

5220 5230 5240 5250 5260 5270 5280 5290

cttactaaaagccagataaaacagtatgcgtattttgogcgctgatttttgoggtataagaatatatactgatatgtatacc

pAG423.seq(1>8396)

5300 5310 5320 5330 5340 5350 5360 5370

cgaagtatgtcaaaaagaggtatgctatgaagcagcgtattacagtgacagttgacagcgcagcgtatcagttgtctcaa

pAG423.seq(1>8396)

5380 5390 5400 5410 5420 5430 5440 5450

ggcatatatgatgtcaatatctccggtctggttaagcacaaccatgcagaatgaagcccgctgctctgctgctgcgaacgct

pAG423.seq(1>8396)

5460 5470 5480 5490 5500 5510 5520 5530

ggaaagcggaaaatcaggaagggtatggctgaggtcgcccggtttattgaaatgaacgggtcttttgctgacgagaaacag

pAG423.seq(1>8396)

5540 5550 5560 5570 5580 5590 5600

gggctggtgaaATGCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGAGTGA

pAG423.seq(1>8396)  
ccdB.seq(1>306)

5610 5620 5630 5640 5650 5660 5670 5680

TATTATTGACACGCCCGGGGCGACGGATGGTGTATCCCGCTGGCCAGTGCACGCTGCTGTGCAGATAAAGTCTCCCGTGAA

pAG423.seq(1>8396)  
ccdB.seq(1>306)

5690 5700 5710 5720 5730 5740 5750 5760

CTTTACCCGCTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCAACCGATATGGCCAGTGTGCCGGTCTCCGTTA

pAG423.seq(1>8396)  
ccdB.seq(1>306)

5770 5780 5790 5800 5810 5820 5830 5840

TCGGGGAAGAAGTGGCTGATCTCAGCCACCGGAAATGACATCAAAAACGCCATTAACTGATGTTCTGGGGAATATA

pAG423.seq(1>8396)  
ccdB.seq(1>306)

5850 5860 5870 5880 5890 5900 5910 5920

Aatgtcaggtcccttatacacagccagtcgtcaggtcgaccatagtgactggatattgttgggttttacagtattatgt

pAG423.seq(1>8396)  
ccdB.seq(1>306)

5930 5940 5950 5960 5970 5980 5990 6000

agtcgtgttttttatgcaaaatctaatttaatatattgatatttatatcattttacggtttctcggttcagctttctgttac

pAG423.seq(1>8396)

6010 6020 6030 6040 6050 6060 6070 6080

aaagtggtgatgggctgcaggaattcgatatcaagcttatcgataccgtcgacctcgagtcagtgtaATTAGTTATGTCA

pAG423.seq(1>8396)  
CYC1-Terminator.seq(1>240)

6090 6100 6110 6120 6130 6140 6150 6160

CGCTTACATTACAGCCCTCCCCCACATCCGCTCTAACCGAAAGGAAGGAGTTAGACAACTGAAGTCTAGGTCCTTA

pAG423.seq(1>8396)  
CYC1-Terminator.seq(1>240)



		6170 6180 6190 6200 6210 6220 6230 6240
		TTTATTTTATAGTTATGTTAGTATTAGAAGCGTTATTATATTCAAATTTTCTTTTTTCTGTACAGACGCGT
pAG423.seq (1>8396)	→	tttattttttatagttatgtagtattagaacggtatttatatttcaatttttctttttttctgtacagacggt
CYC1-Terminator.seq (1>240)	→	TTTATTTTATAGTTATGTTAGTATTAGAAGCGTTATTATATTCAAATTTTCTTTTTTCTGTACAGACGCGT
		6250 6260 6270 6280 6290 6300 6310 6320
		GTACGCATGTAACATTATACTGAAACCTTGCTTGAGAAGGTTTGGGACGCTCGAAGGCTTTAATTTGgggcgggtac
pAG423.seq (1>8396)	→	gtacgcattgaacattatactgaaaccttgcttgagaaggttttgggacgctogaaggctttaatttgggcgggtac
CYC1-Terminator.seq (1>240)	→	GTACGCATGTAACATTATACTGAAACCTTGCTTGAGAAGGTTTGGGACGCTCGAAGGCTTTAATTTG
		6330 6340 6350 6360 6370 6380 6390
		ccaattgcctatagtgagtcgtattacgcgcgtcactggcgtgttttacaacgtcgtgactgggaaacccctgg
pAG423.seq (1>8396)	→	ccaattgcctatagtgagtcgtattacgcgcgtcactggcgtgttttacaacgtcgtgactgggaaacccctgg
		6400 6410 6420 6430 6440 6450 6460 6470
		cggtacccaacttaatgccttgacgacatccccctttgcagctggcgtaatagogaagaggcccgaccgatcgc
pAG423.seq (1>8396)	→	cggtacccaacttaatgccttgacgacatccccctttgcagctggcgtaatagogaagaggcccgaccgatcgc
		6480 6490 6500 6510 6520 6530 6540 6550
		ccttcccaacagttgacgacgtgaatggcgaatggcgacgcgcctgtagcggcgacattaagcggcggtgtgtg
pAG423.seq (1>8396)	→	ccttcccaacagttgacgacgtgaatggcgaatggcgacgcgcctgtagcggcgacattaagcggcggtgtgtg
		6560 6570 6580 6590 6600 6610 6620 6630
		tgggtacgcgcagcgtgacgcgtacacttgccagcgccttagcgcgcctccttgcgtttctcccttcccttctgc
pAG423.seq (1>8396)	→	tgggtacgcgcagcgtgacgcgtacacttgccagcgccttagcgcgcctccttgcgtttctcccttcccttctgc
		6640 6650 6660 6670 6680 6690 6700 6710
		caagttgcgcgctttcccgctcaagctctaaatcgggggctcccttaggggtccgatttagtgctttacggcacctc
pAG423.seq (1>8396)	→	caagttgcgcgctttcccgctcaagctctaaatcgggggctcccttaggggtccgatttagtgctttacggcacctc
		6720 6730 6740 6750 6760 6770 6780 6790
		gaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttgcgcctttgaagt
pAG423.seq (1>8396)	→	gaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttgcgcctttgaagt
		6800 6810 6820 6830 6840 6850 6860 6870
		tggagtccagcttctttaatagtggaactctgttccaaactggaacaacactcaaccctatctcgggtctattcttttga
pAG423.seq (1>8396)	→	tggagtccagcttctttaatagtggaactctgttccaaactggaacaacactcaaccctatctcgggtctattcttttga
		6880 6890 6900 6910 6920 6930 6940 6950
		tttataagggttttgcgatttcggcctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaac
pAG423.seq (1>8396)	→	tttataagggttttgcgatttcggcctatttggttaaaaaatgagctgatttaacaaaaatttaacgcgaattttaac
		6960 6970 6980 6990 7000 7010 7020 7030
		aaaatattaacgtttacaatttccTGATGCGGTATTTCTCTTACGCATCTGTGCGGTATTTACACCGCATagatcc
pAG423.seq (1>8396)	→	aaaatattaacgtttacaatttccTGATGCGGTATTTCTCTTACGCATCTGTGCGGTATTTACACCGCATagatcc
Seq II.seq (1>48)	→	TGATGCGGTATTTCTCTTACGCATCTGTGCGGTATTTACACCGCA
		7040 7050 7060 7070 7080 7090 7100 7110
		gtcgagttcaagagaaaaaaaagaaaaagcaaaaagaaaaaggaagcgcgcctcgttcagaatgacacgtatagaa
pAG423.seq (1>8396)	→	gtcgagttcaagagaaaaaaaagaaaaagcaaaaagaaaaaggaagcgcgcctcgttcagaatgacacgtatagaa
		7120 7130 7140 7150 7160 7170 7180
		tgatgcattaccttgatcattcagtatcatactgttcgtatatacatacttactgacattcataggtatatacatataca
pAG423.seq (1>8396)	→	tgatgcattaccttgatcattcagtatcatactgttcgtatatacatacttactgacattcataggtatatacatataca
		7190 7200 7210 7220 7230 7240 7250 7260
		oatgtatatatatogtatgctgagcgttttaataatoggtgtcactacataagaacaccttgggtggagggaacatcgt
pAG423.seq (1>8396)	→	catgtatatatatogtatgctgagcgttttaataatoggtgtcactacataagaacaccttgggtggagggaacatcgt
His3.seq (1>659)	→	TACATAAGAACACCTTTGGTGGAGGGAACATCGT

Genomic map of the pAG423 plasmid showing the location of the pAG423.seq (1>8396) and HIS3.seq (1>659) regions. The map includes a scale from 7270 to 8290 bp and a detailed view of the 8396 bp pAG423.seq region. The pAG423.seq region is divided into 10 segments, each 8396 bp long. The HIS3.seq region is 659 bp long. The map shows the pAG423.seq (1>8396) and HIS3.seq (1>659) regions in green, and the pAG423.seq (1>8396) and HIS3.seq (1>659) regions in red. The map also shows the pAG423.seq (1>8396) and HIS3.seq (1>659) regions in blue.

Scale: 7270 7280 7290 7300 7310 7320 7330 7340

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7350 7360 7370 7380 7390 7400 7410 7420

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7430 7440 7450 7460 7470 7480 7490 7500

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7510 7520 7530 7540 7550 7560 7570 7580

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7590 7600 7610 7620 7630 7640 7650 7660

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7670 7680 7690 7700 7710 7720 7730 7740

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7750 7760 7770 7780 7790 7800 7810 7820

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7830 7840 7850 7860 7870 7880 7890 7900

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7910 7920 7930 7940 7950 7960 7970

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

7980 7990 8000 8010 8020 8030 8040 8050

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

8060 8070 8080 8090 8100 8110 8120 8130

pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

8140 8150 8160 8170 8180 8190 8200 8210

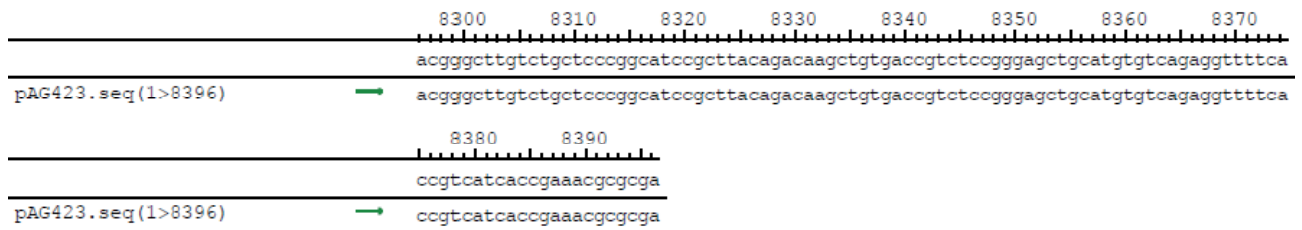
pAG423.seq (1>8396) → →

HIS3.seq (1>659) → →

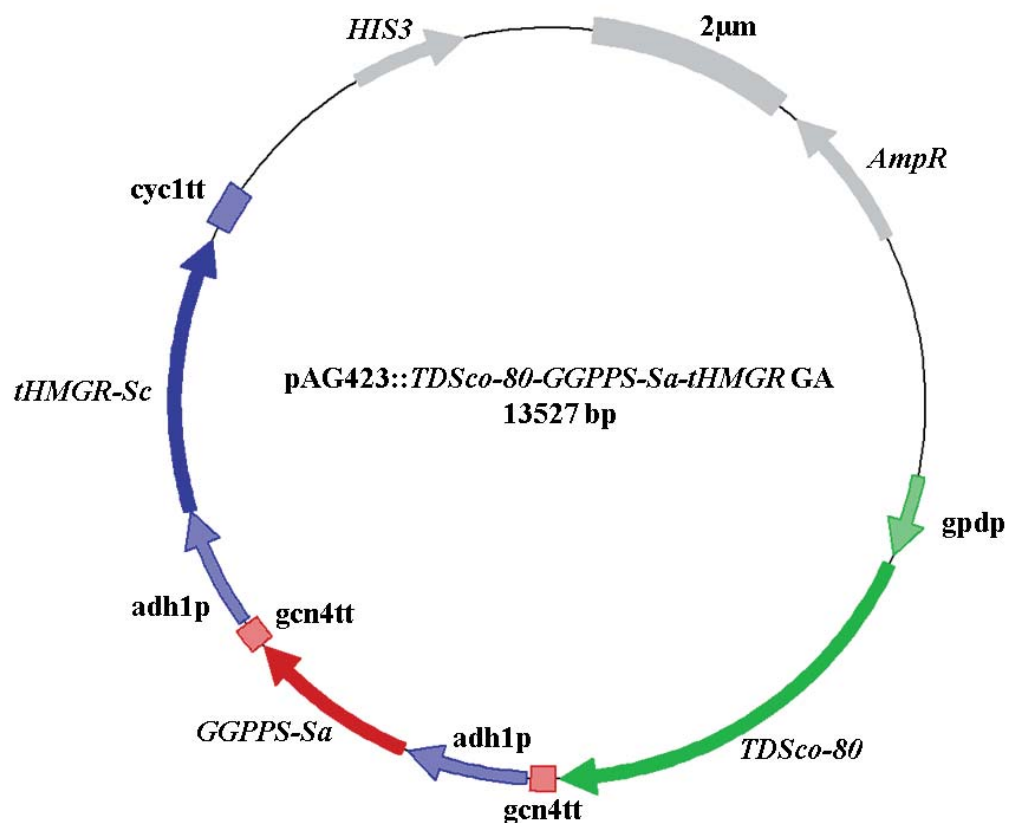
8220 8230 8240 8250 8260 8270 8280 8290

pAG423.seq (1>8396) → →

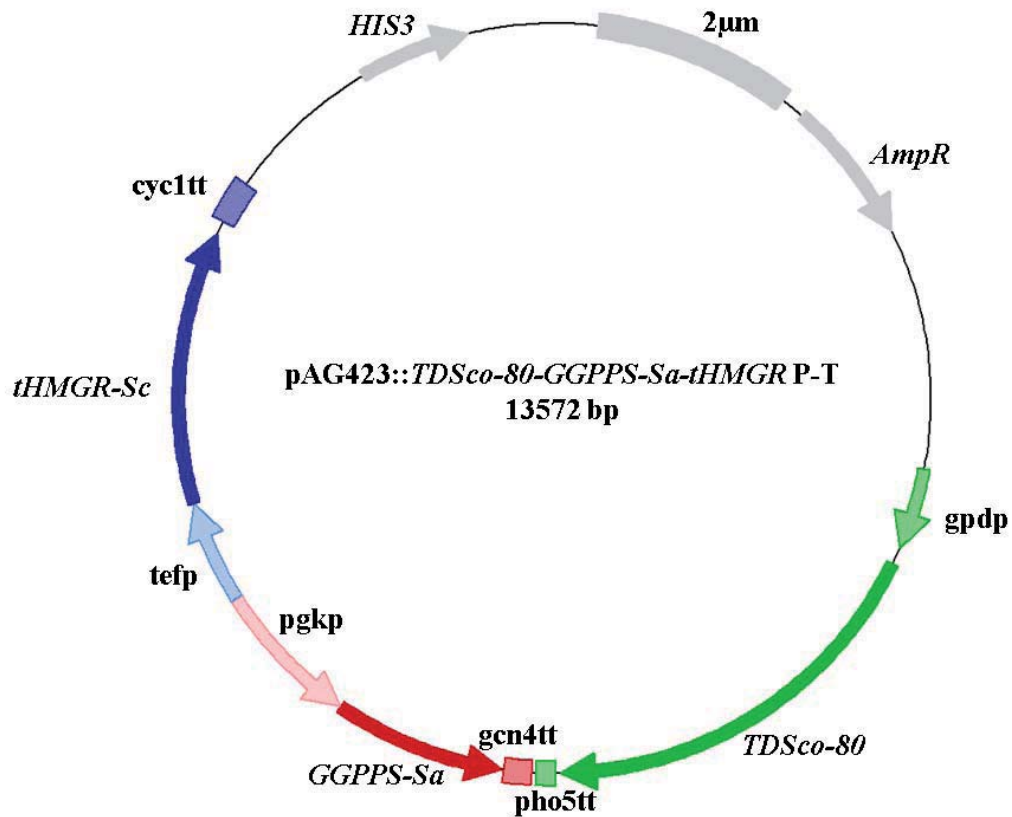
HIS3.seq (1>659) → →



## Vektorkarten

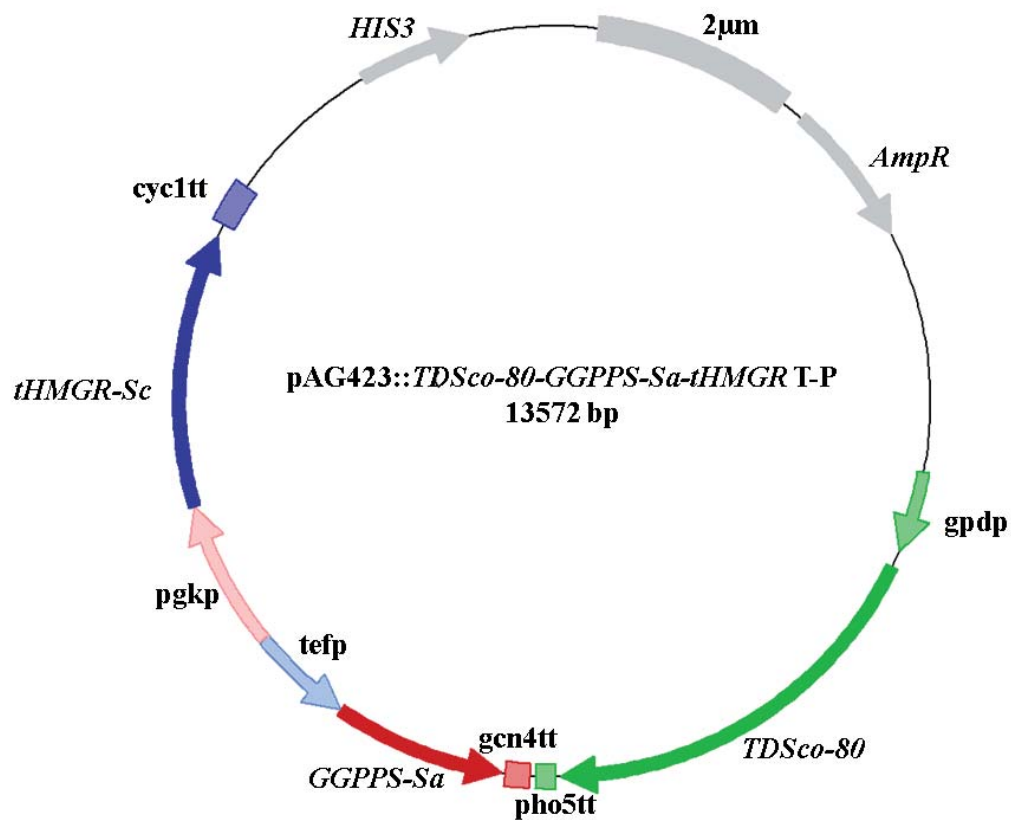


**Fusionskonstrukt pAG423::T-G-t GA.** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind in gleicher Orientierung hintereinander geschaltet. Die Regulationseinheiten zwischen den Genen sind identisch (gcn4tt-adh1p). **HIS3:** Selektionsmarker Histidin; **2 µm:** 2micron; **AmpR:** Ampicillin-Resistenzmarker; **gpdp:** Glukose-6-phosphat DehydrogenasePromotor; **TDSco-80:** codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **GGPPS-Sa:** Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **gcn4tt:** Terminator des Gens für Transkriptionsaktivatorprotein; **adh1p:** Alkoholdehydrogenase-Promotor; **tHMGR-Sc:** verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt:** Cytochrom1-Terminator

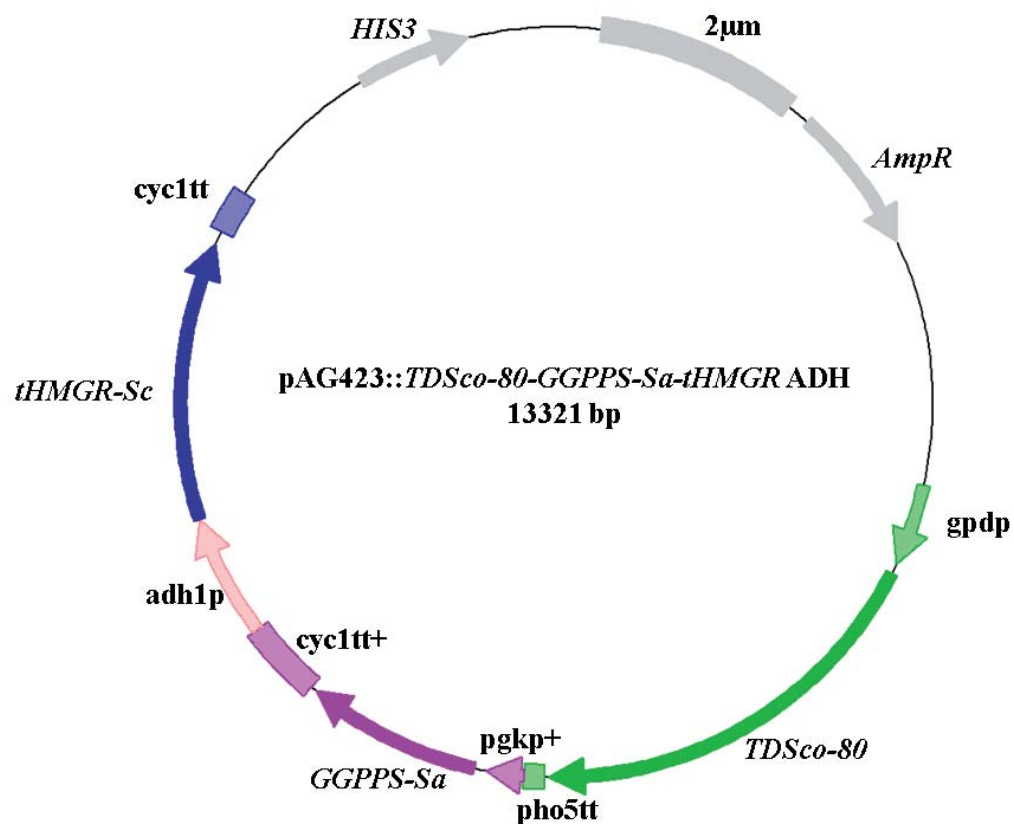


**Fusionskonstrukt pAG423::T-G-t P-T.** Die Gene *TDSco-80* und *GGPPS-Sa* sind konvergent, *GGPPS-Sa* und *tHMGR-Sc* divergenter Orientierung positioniert. ***HIS3***: Selektionsmarker Histidin; **2 µm**: 2micron; ***AmpR***: Ampicillin-Resistenzmarker; ***gdp***: Glukose-6-phosphat Dehydrogenase Promotor; ***TDSco-80***: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; ***pho5tt***: Saure Phosphatase-Terminator; ***gcn4tt***: Terminator des Gens für Transkriptionsaktivatorprotein; ***GGPPS-Sa***: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; ***pgkp***: Phosphoglyceratkinasepromotor; ***tefp***: Promotor des Transkriptionselongationsfaktorgens; ***adh1p***: Alkoholdehydrogenase-Promotor; ***tHMGR-Sc***: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; ***cyc1tt***: Cytochrom1-Terminator

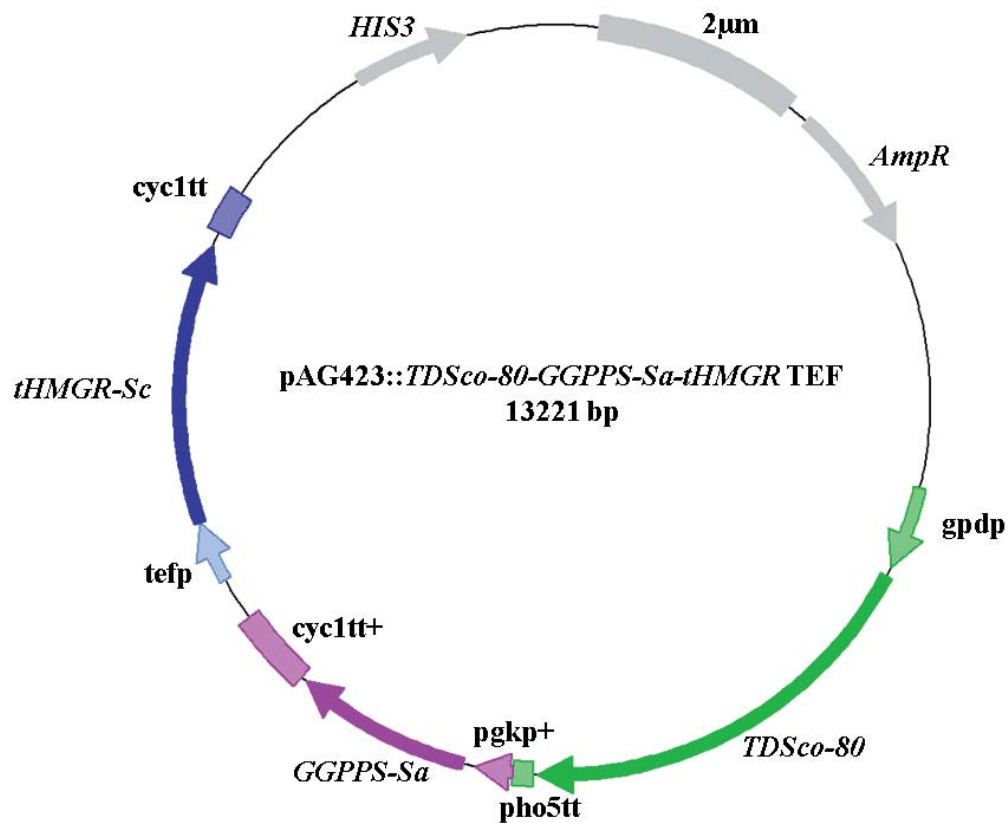




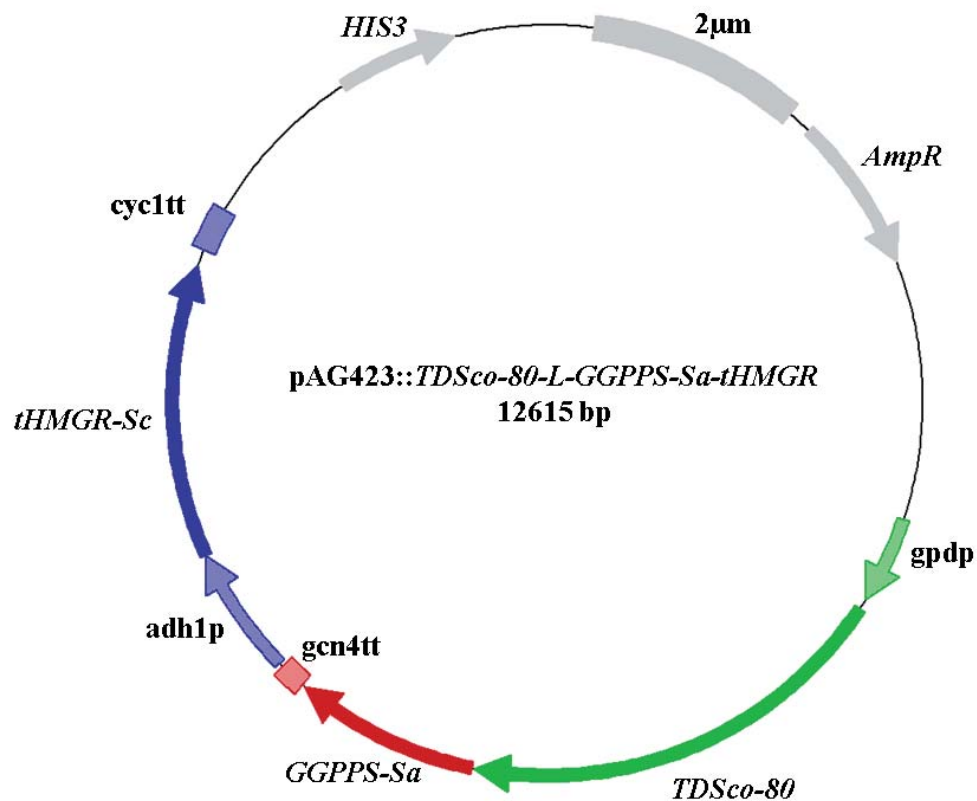
**Fusionskonstrukt pAG423::T-G-t T-P.** Die Gene *TDSco-80* und *GGPPS-Sa* sind konvergent, *GGPPS-Sa* und *tHMGR-Sc* divergenter Orientierung positioniert. **HIS3**: Selektionsmarker Histidin; **2 µm**: 2micron; **AmpR**: Ampicillin-Resistenzmarker; **gdpd**: Glukose-6-phosphat Dehydrogenase Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **pho5tt**: Saure Phosphatase-Terminator; **gcn4tt**: Terminator des Gens für Transkriptionsaktivatorprotein; **GGPPS-Sa**: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **pgkp**: Phosphoglyceratkinasepromotor; **tefp**: Promotor des Transkriptionselongationsfaktorgens; **adh1p**: Alkoholdehydrogenase-Promotor; **tHMGR-Sc**: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt**: Cytochrom1-Terminator



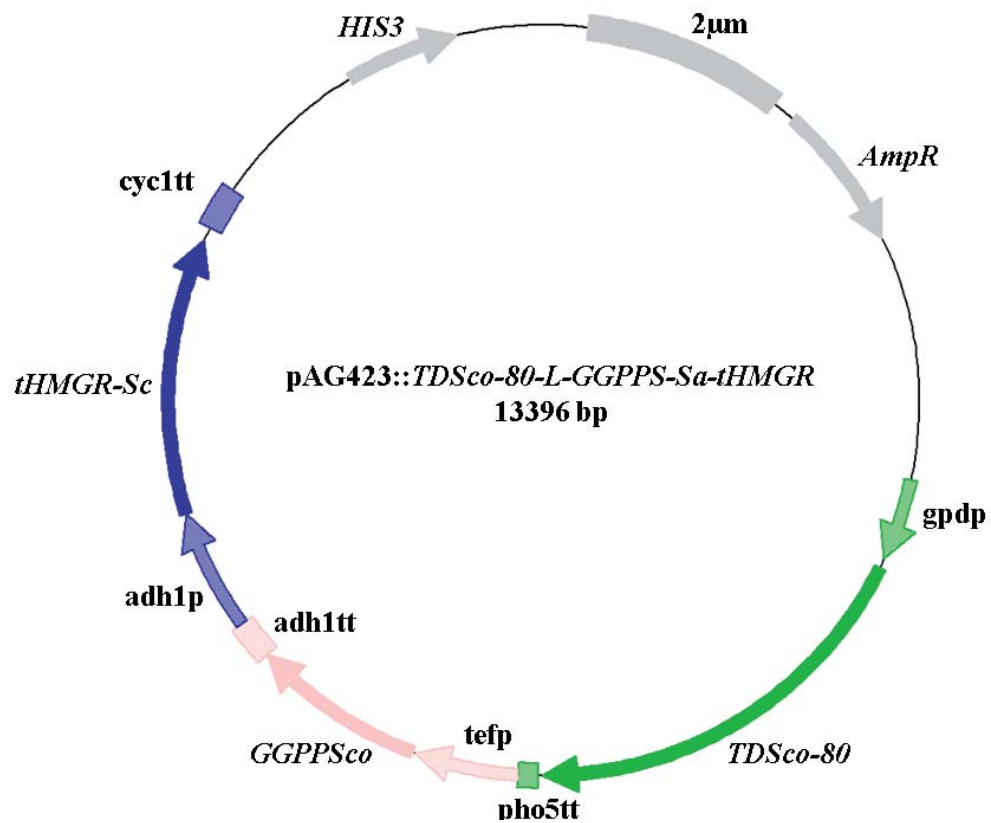
**Fusionskonstrukt pAG423::T-G-t ADH.** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind im Tandem orientiert. Die Einheit *pgkp<sup>+</sup>-GGPPS-Sa-cyc1tt<sup>+</sup>* wurde aus dem Einzelvektor *pVV200::GGPPS-Sa* amplifiziert und anstelle von *gcn4tt-GGPPS-Sa-tefp* in *pAG423::T-G-t* T-P eingesetzt. **HIS3**: Selektionsmarker Histidin; **2 µm**: 2micron; **AmpR**: Ampicillin-Resistenzmarker; **gdp**: Glukose-6-phosphat Dehydrogenase Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **pho5tt**: Saure Phosphatase-Terminator; **pgkp<sup>+</sup>**: Phosphoglyceratkinasepromotor aus *pVV200*; **GGPPS-Sa**: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **cyc1tt<sup>+</sup>**: Cytochrom1-Terminator aus *pVV200*; **adh1p**: Alkoholdehydrogenase-Promotor; **tHMGR-Sc**: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt**: Cytochrom1-Terminator



**Fusionskonstrukt pAG423::T-G-t ADH.** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind im Tandem orientiert. Die Einheit *pgkp+*-*GGPPS-Sa*-*cyc1tt+* wurde aus dem Einzelvektor pVV200::*GGPPS-Sa* amplifiziert und anstelle von *gen4tt*-*GGPPS-Sa*-*tefp* in pAG423::*T-G-t* P-T eingesetzt. **HIS3**: Selektionsmarker Histidin; **2 μm**: 2micron; **AmpR**: Ampicillin-Resistenzmarker; **gdp**: Glukose-6-phosphat Dehydrogenase Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **pho5tt**: Saure Phosphatase-Terminator; **pgkp+**: Phosphoglyceratkinasepromotor aus pVV200; **GGPPS-Sa**: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **cyc1tt+**: Cytochrom1-Terminator aus pVV200; **tefp**: Promotor des Transkriptionselongationsfaktorgens; **tHMGR-Sc**: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt**: Cytochrom1-Terminator



**Fusionskonstrukt pAG423::T-G-t L.** Die Gene *TDSco-80*, *GGPPS-Sa* und *tHMGR-Sc* sind im Tandem orientiert. *TDSco-80* und *GGPPS-Sa* sind über Gly-Gly-Gly-Ser miteinander verbunden und kodieren für ein Fusionsprotein. **HIS3**: Selektionsmarker Histidin; **2 µm**: 2micron; **AmpR**: Ampicillin-Resistenzmarker; **gdp**: Glukose-6-phosphat Dehydrogenase Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthese; **GGPPS-Sa**: Geranylgeranyldiphosphatsynthasegen aus *Sulfolobus acidocaldarius*; **gen4tt**: Terminator des Gens für Transkriptionsaktivatorprotein; **adh1p**: Alkoholdehydrogenase-Promotor; **tHMGR-Sc**: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt**: Cytochrom1-Terminator



**Fusionskonstrukt pAG423::T-G-t co.** Die Gene *TDSco-80*, *GGPPSco* und *tHMGR-Sc* sind im Tandem orientiert. *GGPPSco* inklusive Promotor und Terminator wurden synthetisch hergestellt. **HIS3**: Selektionsmarker Histidin; **2 µm**: 2micron; **AmpR**: Ampicillin-Resistenzmarker; **gdpd**: Glukose-6-phosphat Dehydrogenase Promotor; **TDSco-80**: codon-optimiertes Gen für um 80 Aminosäuren verkürzte Taxadiensynthase; **pho5tt**: Saure Phosphatase-Terminator; **tefp**: Promotor des Transkriptionselongationsfaktorgens; **GGPPSco**: codon-optimiertes Geranylgeranyldiphosphat-synthasegen aus *Sulfolobus acidocaldarius*; **adh1tt**: Alkoholdehydrogenase-Terminator; **adh1p**: Alkoholdehydrogenasepromotor; **tHMGR-Sc**: verkürztes HMG CoA-Reduktasegen aus *Saccharomyces cerevisiae*; **cyc1tt**: Cytochrom1-Terminator

# Lebenslauf

## Persönliche Daten

Geburtsdatum	14.12.1977
Geburtsort	Münster
Familienstand	ledig
Nationalität	deutsch, schweiz

## Anstellungen

Okt. 2008 – Feb. 2009	Lehrtätigkeit (Praktikumsbetreuung Pflanzenbiotechnologie) Institut für Nano- und Biotechnologie Fachhochschule Aachen, Campus Jülich
Sep. 2005 – Dez. 2010	<u>PROMOTIONSARBEIT</u> : “Metabolic engineering der Taxolbiosynthese in <i>Saccharomyces cerevisiae</i> ” im Arbeitskreis Prof. Dr. W.-D. Fessner; extern angefertigt am Fraunhofer Institut für Molekularbiologie und angewandte Ökologie, Aachen

## Studium

WS 1998 – SS 2005	Diplom-Biologie Westfälische Wilhelms-Universität zu Münster
SS 2004 – SS 2005	<u>DIPLOMARBEIT</u> : „Untersuchungen an putativen Genen zum Cyanophycinstoffwechsel in <i>Ralstonia eutropha</i> H16, bzw. deren Genprodukten“. Institut für molekulare Mikrobiologie und Biotechnologie der WWU zu Münster

## Praktikum

Feb. 2002 – Okt. 2002	Syngenta Crop Protection Research Biology, Stein-Säckingen, CH
-----------------------	---

## Schulausbildung

Aug. 1988 – Jul. 1998	Immanuel-Kant-Gymnasium, Münster
Jul. 1994 – Jun. 1995	Auslandsaufenthalt Canyon Del Oro Highschool, Arizona, USA

Pia Dahm, Aachen, 14.10.2010



Pia Dahm  
Viktoriastr. 48  
52066 Aachen  
Tel.: 0163 – 4705582  
e-mail: dahm@molbiotech.rwth-aachen.de

Aachen, 14.10.2010

### **Eidesstattliche Erklärung**

Ich erkläre hiermit an Eides Statt, dass ich meine Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

Pia Dahm

Pia Dahm  
Viktoriastr. 48  
52066 Aachen  
Tel.: 0163 – 4705582  
e-mail: dahm@molbiotech.rwth-aachen.de

Aachen, 14.10.2010

### **Erklärung**

Ich erkläre hiermit, noch keinen Promotionsversuch unternommen zu haben.

Pia Dahm

## **Danksagung**

Ich danke

Prof. Dr. W.-D. Fessner für die Überlassung des spannenden Themas und sein organisatorisches Engagement

Prof. Dr. H. Kolmar für die Übernahme des Korreferats

Dr. S. Jennewein für die wissenschaftliche Betreuung und Unterstützung vor Ort

Dipl.-Ing. B. Engels und Dipl.-Ing. S. Scholz für die Mithilfe im Rahmen ihrer Diplomarbeiten

Dipl.-Ing U. Heinig für seinen fachlichen Beistand, besonders in den letzten Monaten

Dipl.-Biologen O. Mahmoud und B. Otte für die angenehme Stimmung und den regen Austausch im Büro

den Arbeitskollegen der Weißen Biotechnologie-, sowie der Pflanzenbiotechnologie- und IPP-Gruppe des Fraunhofer IME

meinem Freund Lars für seine Zuversicht

meinen Eltern für ihre offenen Ohren in wirklich jeder Situation